ODICAL ROOM
RAL LIBRARY
IT. OF MICH.

DEC 21 1923

Journal of Educational Research

OFFICIAL ORGAN OF THE
EDUCATIONAL RESEARCH ASSOCIATION

B. R. BUCKINGHAM, EDITOR

ASSOCIATE EDITORS

E. J. ASHBAUGH — W. W. CHARTERS — GEORGE D. STRAYER
GUY M. WHIPPLE — S. A. COURTIS — LEWIS M. TERMAN
HARRY A. GREENE

VOL. VIII DECEMBER, 1923 No. 5

CONTENTS

		Page
PUTTING TEST SCORES IN A FORM TO FACILITATE COMPARISONS	T. V. Goodrich	377
THE PRINCIPLE OF THE SINGLE VARIABLE IN A SPEED OF READING CROSS-OUT TEST	J. Crosby Chapman and S. Cook	389
THE RHODE ISLAND INTELLIGENCE TEST	Grace E. Bird	397
THE COMPARATIVE RANKING OF HIGH SCHOOLS BASED UPON RECORDS AT THE UNIVERSITY OF WASHINGTON	Alexander C. Roberts	404
PROPOSALS FOR STANDARDIZING MEASUREMENT IN EDUCATION: I. THE VARIABILITY UNIT	Harry S. Will	414
THE EFFECT OF ATTENDANCE UPON SCHOOL ACHIEVEMENT	C. W. Odell	422
A GRAPHIC RATING SCALE FOR TEACHERS	Max Freyd	433
THE OTIS CORRELATION CHART	Arthur S. Otis	440
EDITORIALS		449
REVIEWS AND ABSTRACTS		453
NEWS ITEMS AND COMMUNICATIONS		460

Entered as second-class matter at the Post Office at Bloomington, Illinois under the act of March 3, 1879.

Journal of Educational Research

Copyright 1923, by
Public School Publishing Co.

Official Organ of the Educational Research Association

Editor-in-Chief, B. R. Buckingham
Director of Bureau of Educational Research, Ohio State University, Columbus, Ohio

Associate Editors

E. J. ASHBAUGH, *Ohio State University, Columbus.*

W. W. CHARTERS, *Carnegie Inst. Tech., Pittsburgh.*

S. A. COURTIS, *City Schools, Detroit, Mich.*

HARRY A. GREENE, *Ex-Officio, Secretary Educational Research Association, State University of Iowa, Iowa City.*

GEORGE D. STRAYER, *Teachers College, Columbia University, New York City.*

LEWIS M. TERMAN, *Leland Stanford Junior University, Stanford University, California*

GUY M. WHIPPLE, *University of Michigan, Ann Arbor, Michigan.*

Contributing Editors

VIRGIL E. DICKSON, *Director, Bureau of Research and Guidance, Public Schools, Oakland and Berkeley, California.*

CHARLES FORDYCE, *Director of Bureau of Educational Measurements, University of Nebraska, Lincoln, Nebr.*

M. E. HAGGERTY, *Dean of the College of Education, University of Minnesota, Minneapolis.*

V. A. C. HENMON, *Director, School of Education, University of Wisconsin, Madison.*

ARTHUR W. KALLOM, *Assistant Director, Department of Educational Investigation and Measurement, Public Schools, Boston, Mass.*

F. J. KELLY, *Dean of University Administration, University of Kansas, Lawrence, Kans.*

ROBT. H. LANE, *Assistant Superintendent of Schools Los Angeles, Cal.*

WILLIAM A. McCALL, *Assistant Professor of Education, Teachers' College, Columbia University.*

GEORGE W. MELCHER, *Director, Bureau of Research and Efficiency, Public Schools, Kansas City, Missouri.*

EUGENE A. NIFENECKER, *Assistant Director, Bureau of Reference, Research and Statistics, Public Schools, New York City.*

JOSEPH P. O'HERN, *Assistant Superintendent of Schools and Director of Efficiency Bureau, Public Schools, Rochester, New York.*

LEON O. SMITH, *Assistant Superintendent (in charge of Educational Research) Public Schools, Omaha, Nebraska.*

W. W. THEISEN, *Assistant Superintendent of Schools Milwaukee, Wis.*

CLIFFORD WOODY, *Director Bureau of Tests and Measurements, University of Michigan, Ann Arbor, Mich.*

The JOURNAL OF EDUCATIONAL RESEARCH is published at Bloomington, Illinois, monthly, except July and August by the Public School Publishing Co. The subscription price is $4.00 $4.20 in Canada or Mexico; and $4.40 in other foreign countries.

The editorial office is at Columbus, Ohio. Address all communications on editorial matters to B. R. Buckingham, Director of Bureau of Educational Research, The Ohio State University, Columbus, Ohio.

Subscriptions and all communications about reprints or other business matters should be addressed to the Public School Publishing Company at Bloomington, Illinois.

JOURNAL *of* EDUCATIONAL RESEARCH

Volume VIII DECEMBER, 1923 Number 5

PUTTING TEST SCORES IN A FORM TO FACILITATE COMPARISONS

T. V. GOODRICH
Bureau of Research, Lincoln, Nebraska

A serious obstacle in the way of securing the greatest possible benefit from the use of standard tests lies in the fact that the scores in different tests have no common significance and that direct comparisons cannot easily be made between the achievement of the same pupils or classes in different subjects. A score of twelve, for example, represents sixth-grade ability in the Trabue Language scale, fifth-grade ability in the Monroe Reading test, and third-grade ability in the Charters Pronoun test. It is above the upper limit of the Buckingham scale for problems in arithmetic and below the lower limit of the Ayres Penmanship scale. In this respect "standard" tests are not standardized. It is obvious that such a condition makes the interpretation of scores difficult and thereby somewhat impairs the usefulness of the tests.

Among the difficulties which confront the teacher or school administrator who desires to make standard tests yield their greatest measure of usefulness the three following are serious and grow out of the situation just stated: (1) Constant reference to a table of norms is necessary in interpreting results. The only alternative to this is to memorize a mass of figures which have no other significance than their application to a particular scale. (2) An explanation of the significance of scores, not at all easy in most cases, is necessary every time an attempt is made to inform a parent where his child stands, or to explain to an interested citizen what a class or school has accomplished. (3)

Comparisons are difficult to make, except with reference to a single ability as measured by some one particular test. Quantitative comparisons involving more than this require a computation of more or less complexity. It naturally follows that few are made. Graphic comparisons, the most effective form to use in the great majority of cases, cannot be made until the scores, by some method, have been stated in terms of a common unit.

The possibilities of graphic representation of test results was the chief consideration which led the Bureau of Research in the Lincoln public schools to make a study of the methods by which scores could be reduced to a common basis and to adopt the method described in this paper.

The conditions giving rise to the problem were as follows: During the year 1920-21 six scholarship tests were given throughout the entire system in grades three to eight; each was repeated in parallel form after an interval of about four months. Intelligence tests and several other scholarship tests were given on a smaller scale. It was the practice of the Bureau to include in the reports made to the schools after each test two graphs; one showing the curve of the median scores for all schools in comparison with the norms published by the author; and the other showing the same thing for the particular school to which the report was made. The forms for these graphs were mimeographed on paper of ordinary note-book size. Since each scale or test used required a special form and some tests one for each of several different phases or parts, the preparation of forms on which to make these graphs was no small task. Their value was somewhat limited, too, by their small size, although they were as large as could be conveniently made with the mimeograph. Despite their small size, the graphs formed the part of these reports which was most used. Both teachers and pupils studied them with interest and obtained a much clearer view of the situation than they would have obtained from the tabular and descriptive part alone.

It was apparent that if a single form could be devised, on which the results of all tests could be represented, this form

could be printed at a somewhat smaller cost and used with a great saving of time and effort. The charts could be made larger, large enough to be easily seen across a classroom and their value increased in this way. Most important of all, a single form would make possible the comparisons of scores in different subjects and different phases of the same subject. If it were adapted to use with intelligence tests as well as educational tests, it would even provide a simple method for measuring the efficiency of instruction. The plan adopted has other advantages which will be mentioned, but these were the chief considerations that led to its adoption.

Several possible methods were considered. The use of batteries of tests constructed for the special purpose of yielding comparable scores in different subjects, such as the Illinois or Pintner tests, offered a plan already worked out, but a plan was wanted which might be used with any good test. The use of the percentile rank method was considered, but this seemed impracticable because most authors of tests have not published percentile tables, or distribution tables from which percentile norms may be computed. The use of the standard deviation as the unit with which to construct a common scale had a similar disadvantage. The necessary data were seldom available. Moreover, the body of teachers would have little idea of the significance of such a unit. The method of recording scores in terms of percents of the standard score, along lines recently described by Superintendent Brooks and Professor Lindsay, was suggested; but the suggestion was not followed up because direct comparison of progress from grade to grade seemed impossible by this method. The method finally adopted was essentially the one proposed by Doctor Buckingham in the December, 1920, issue of the JOURNAL OF EDUCATIONAL RESEARCH.

The reduction of scores to a common scale based on grade norms in which the unit is a normal year's progress, is the basis of this plan. A number on the scale represents the norm for pupils just entering the grade which we designate by that number. A score of four, for example, represents performance to

be expected of a IV-B (low) class just beginning the semester. Fractional parts of a year's progress are represented by fractional scores. A score of 4.5 represents performance expected of a IV-B class just finishing the semester or a IV-A class just beginning the semester. Scores are changed from the original scale into the common scale by means of tables. The method of constructing a table for any given test is to equate the norms published by the author with the numbers used to describe the grade to which they apply, and to add one-tenth of the difference between the norm for a given grade and that of the next higher grade for each one-tenth of a unit on the new scale. For example, a score of 9.0 in the Woody Addition scale is the norm for a third-grade class in November or December. This is called 3.3 on the new scale. One-tenth the difference between 11 and 9, which is 0.2, represents one-tenth of a year's progress at this point and the new scale is constructed as follows:

9.0	on	the	Woody	scale	equals	3.3	on	the	new	scale
9.2	"	"	"	"	"	3.4	"	"	"	"
9.4	"	"	"	"	"	3.5	"	"	"	"
9.6	"	"	"	"	"	3.6	"	"	"	"
9.8	"	"	"	"	"	3.7	"	"	"	"

No very satisfactory method has been devised for handling scores below the third-grade or above the eighth-grade norms. In practice we have had persons familiar with the type of work measured by this scale estimate what a second- and a tenth-grade pupil should do. The scale is extended to 2.0 and to 10.0 on the basis of these estimates. No scores have been given below 2.0 or above 10.0.

Some difficulties have also been found in preparing tables for tests such as the Monroe Reading test which are divided into parts that do not dovetail. The preparation and publication of authoritative tables of this sort for all well-known and well-standardized tests would be a worth-while contribution if school people generally were to find this plan as useful as we have found it.

The advantages in the use of this scale lie along these lines:

1. It enables us to state measurements by different scales in common terms.
2. The score itself has a meaning—"three" means third grade and "four" means fourth grade. No teacher will ever forget this meaning or have difficulty in explaining its significance to a pupil or a parent.
3. Differences between scores made by different pupils or groups, or by the same pupil or group at different times, are stated in a unit which is significant and easily understood, viz: the progress normally made in a year or a fraction of a year.
4. Scores may be averaged and significant measures may be obtained which are not possible with the scales in their original forms. For example, the mean of the scores in the four operations in arithmetic may be taken as the arithmetic score, the mean of the verb and pronoun scores as a language score, the mean of the scores in all subjects, weighted if desired, as a measure of general scholarship. The mean variation from the norms of all class scores in a given subject, signs being considered, may be taken as the school's index of accomplishment in that subject. "General scholarship" scores for different grades treated in the same way give an index of general scholarship.
5. Comparison of intelligence and achievement scores may be made and a pupil's achievement considered with reference to his ability.
6. Results in all tests and averages as explained in (4) may be plotted on a single form. This not only simplifies the problem of making graphs, but offers many opportunities for graphic comparison wholly impossible where the scales are used in their original forms.

Illustrations are given of charts used to show the following relationships:

Figure 1. Semester's progress in a single subject—one school. The advantages of this chart over the same form

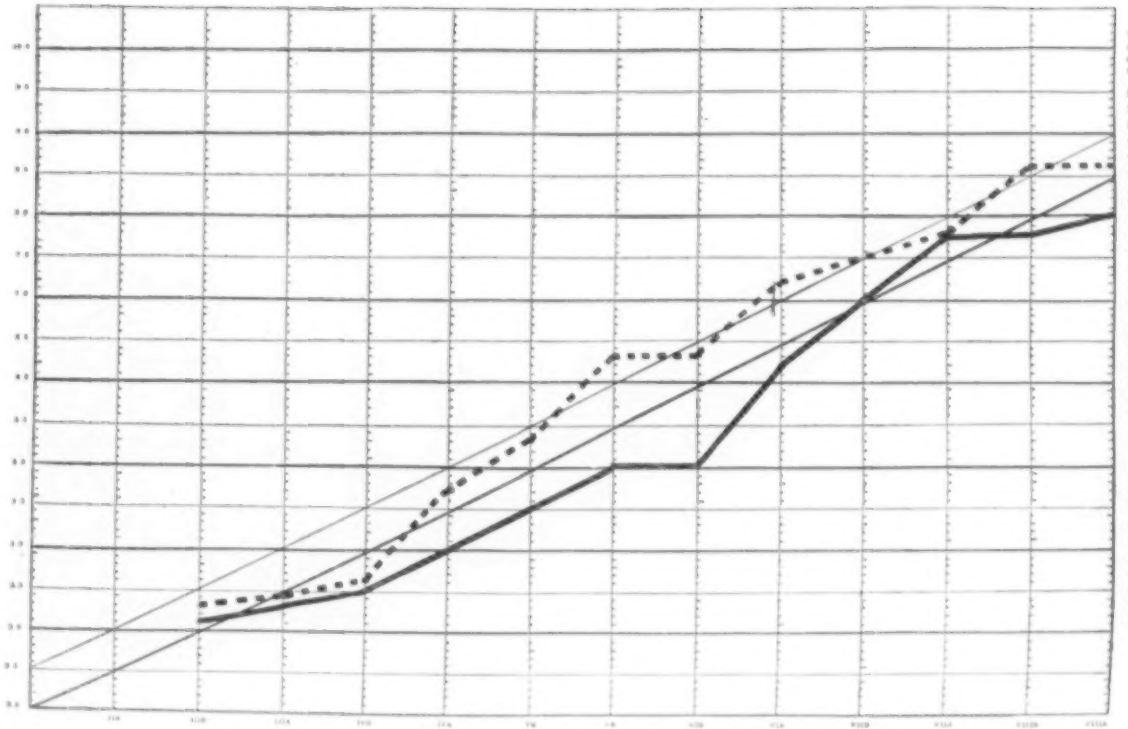


FIGURE 1. CLASS SCORES IN ARITHMETIC PROBLEMS IN TERMS OF GRADE STANDARDS. CLINTON SCHOOL, BUCKINGHAM TEST. FULL LINE FOR FEBRUARY TESTING. DASHED LINE FOR MAY TESTING

FIGURE 1. CLASS SCORES IN ARITHMETIC PROBLEMS IN TERMS OF GRADE STANDARDS. CLINTON SCHOOL, BUCKINGHAM TEST. FULL LINE FOR FEBRUARY TESTING. DASHED LINE FOR MAY TESTING

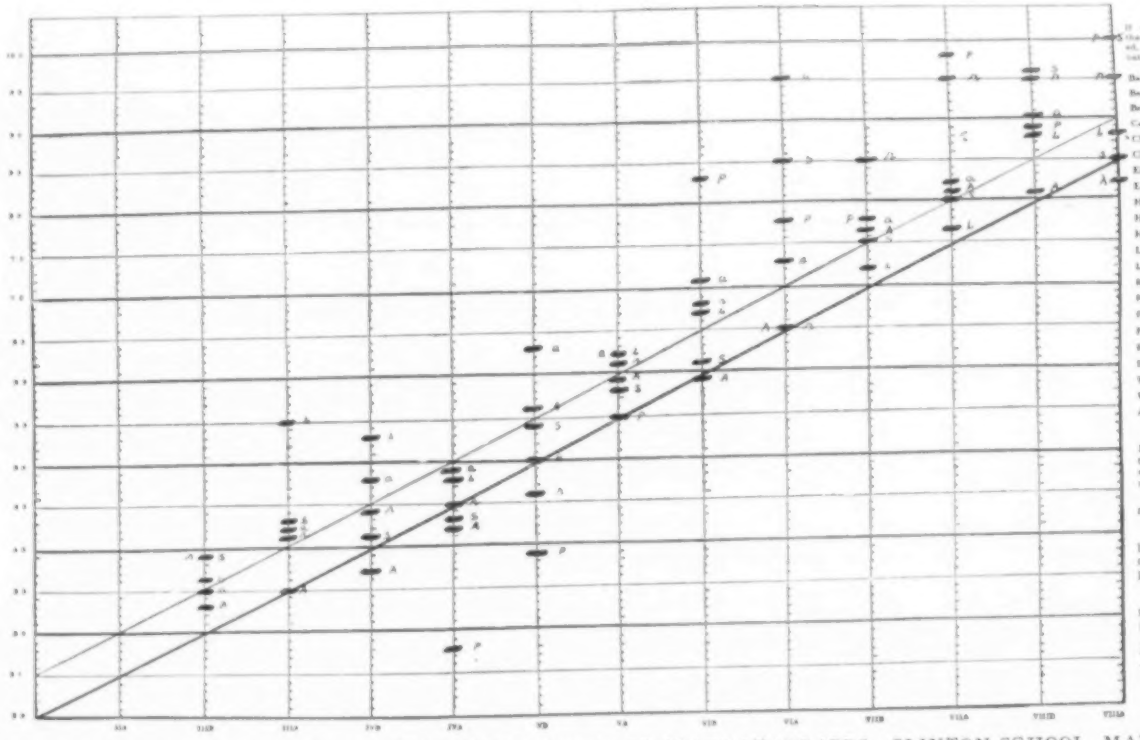


FIGURE 2. CLASS SCORES IN SIX SUBJECTS IN TERMS OF GRADE STANDARDS. CLINTON SCHOOL, MAY, 1921

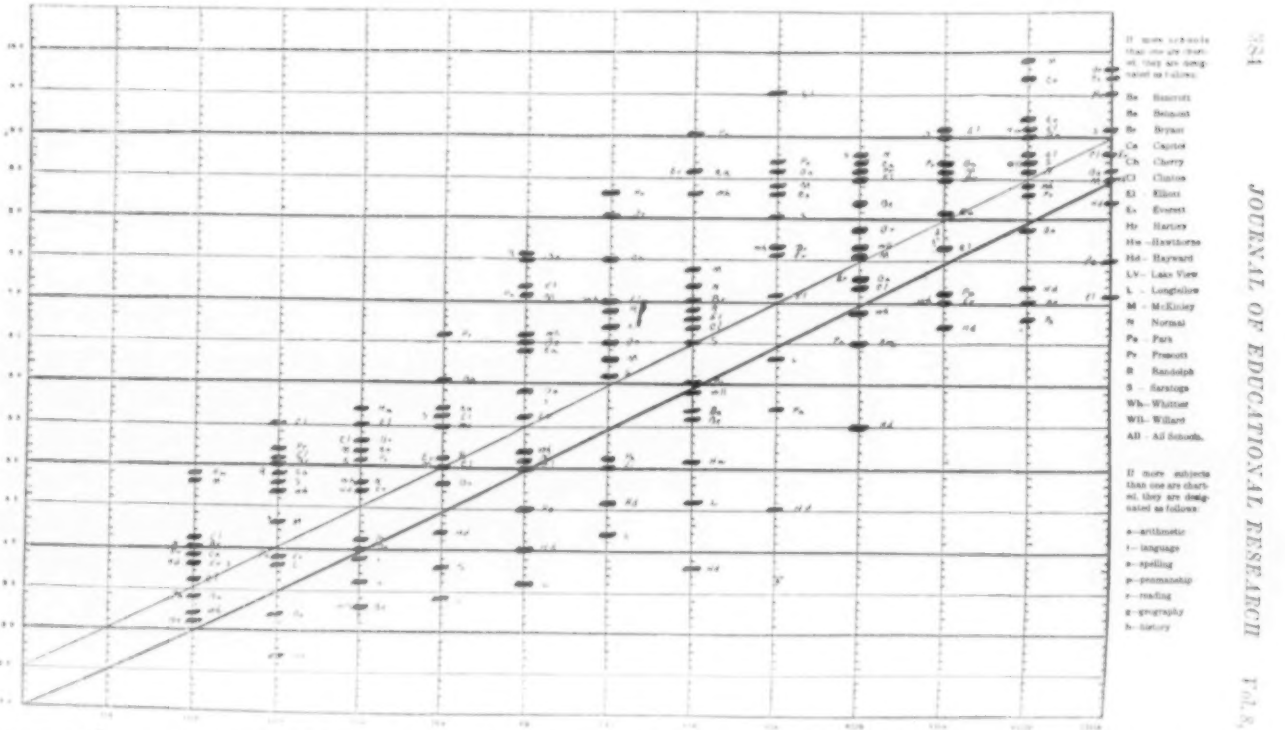


FIGURE 3. CLASS SCORES IN LANGUAGE IN TERMS OF GRADE STANDARDS, ALL SCHOOLS, TRABUE TEST, JANUARY, 1921

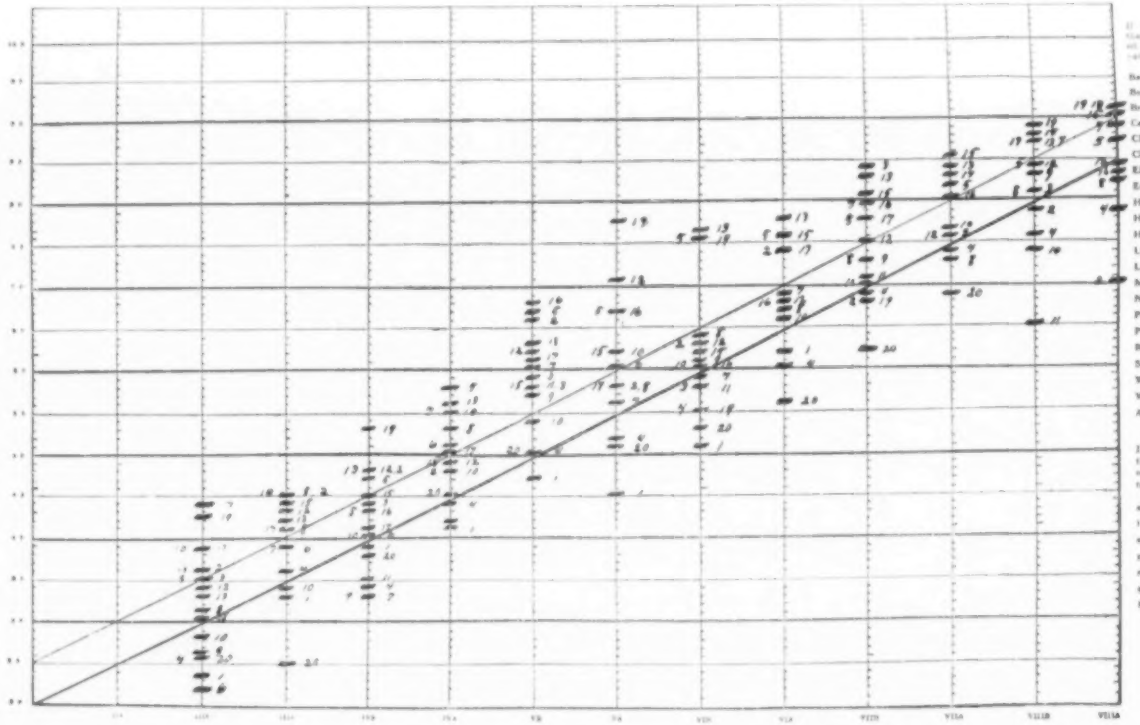


FIGURE 4. GENERAL SCHOLARSHIP IN TERMS OF GRADE STANDARDS EIGHTEEN SCHOOLS, JANUARY AND MAY, 1921

plotted from the original scores lie in the directness with which the amount of progress may be read. The V-A classes, for example, made in the four months as much progress as a class ordinarily makes in 1.3 years. The comparison of any score with the norm is likewise direct and may be stated in the same terms. The IV-A class, for example, in the February test (the beginning of the semester) was one-half year below the norm.

Figure 2. Comparisons of scores in several subjects—one school. This chart is primarily for the use of the building principal, to assist her to determine what subjects need special emphasis; what teachers seem to need her help and what teachers may probably have something in the way of method to contribute to the other teachers. The arithmetic score here represented is the mean of scores in the four operations; the reading score is the mean of the rate and comprehension scores.

Figure 3. Bird's eye view of the achievement of every class in a school system in a single subject. This chart is for the use of the superintendent of schools or for the supervisor of special subjects or grades. It enables him to know where he can spend his time to best advantage. The language scores of all classes in the eighteen schools are here represented.

Figure 4. Comparison of general scholarship in all schools in a system. The scores here represented are weighted averages of the scores in the six different subjects and represent "general scholarship." In actual practice the abbreviations of the names of the schools are used as in Figure 3. The numbers here used do not correspond with order of schools given in the right-hand margin.

These illustrations do not exhaust the possibilities of making comparisons by this method, but should be sufficient to demonstrate its usefulness.

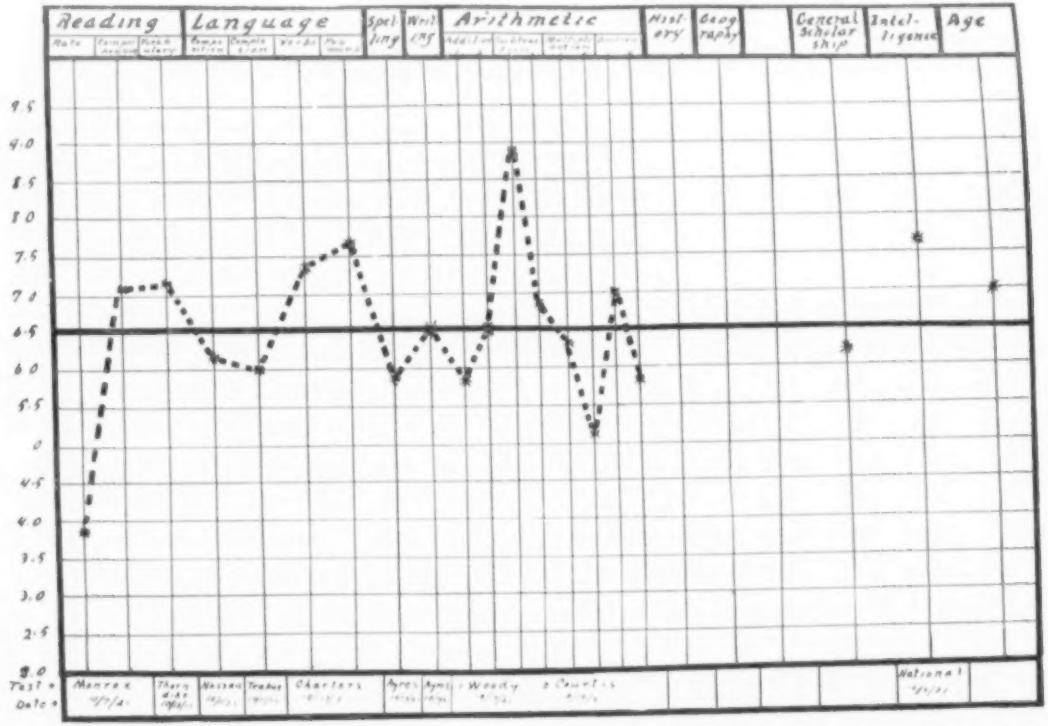


FIGURE 5. STANDARD TEST SCORES OF JOHN BROWN IN TERMS OF GRADE STANDARDS. CAPITOL SCHOOL, GRADE VI-A. DATE OF BIRTH, MARCH 8, 1909

The throwing of individual scores into terms of grade norms assists the teacher in making individual diagnosis. The method, however, is difficult to apply when the unit in the original scale is large, as in the Ayres penmanship scale. In this scale a score of 50 may represent fourth-, fifth-, or sixth-grade work. If a pupil within any one of these grades submits a sample which receives a mark of fifty, it is necessary to attempt to revise the mark with reference to fractional parts of the class interval or to assume that the quality represented is the norm for his grade. In most cases fractional amounts of a year's progress smaller than one-half cannot be represented by individual scores. In spite of these limitations the method is very useful in handling individual scores. It has been used in connection with a complete reclassification of the pupils in one of the Lincoln schools. Figure 5 is given to illustrate the use of the method in interpreting the scores of individual pupils. This form has not been adopted for general use in the Lincoln schools. We hope that this or some similar form will prove to have sufficient merit to justify its adoption.

Summary

1. The usefulness of standardized tests is impaired by the difficulty of interpreting a variety of scores which have no common significance.
2. The reduction of scores in the original scales to terms of a scale based on grade norms in which the unit is a year's progress offers a method of obviating this difficulty.
3. Such a scale is easy to understand and use. With it significant averages can be computed and comparisons, especially in graphic form, made which have hitherto been impossible to make without elaborate computations.
4. The method is useful in diagnosing individual needs and in classifying pupils.
5. Its usefulness would be extended by the compilation and publication of authoritative tables for reducing scores in the most useful scales to this common scale.

THE PRINCIPLE OF THE SINGLE VARIABLE IN A SPEED OF READING CROSS-OUT TEST

J. CROSBY CHAPMAN, *Yale University*
AND
S. COOK, *Rutgers College*

In *The Measurement of Silent Reading* Mrs. Burgess[1] makes a plea that the principle of the single variable be employed in constructing tests if later the scores are to be interpreted in the usual quantitative manner. If we wish to speak of one achievement being double another, we must make doubtful assumptions or else have approximations to single variable scales. But is it necessary to speak of one achievement being double another in an absolute sense? When considering such a complex function as silent reading ability, in what particular school situations is this extremely advantageous? There is much to be said for Buckingham's suggestion that the unit in which the school man thinks is not absolute but rather a function of the improvement made annually, or of the average achievements of age or grade levels.

Before criticising Mrs. Burgess' test, it will be useful, however, to consider at some length the range of application of the principle of the single variable, since her presentation would seem to indicate that any scale with two variables is unscientific.

Whenever we measure a complex ability, the large number of minor abilities which contribute to the complex ability are producing their effect on the score. Unless we are able to make the material in the elements comprising the test equal in difficulty, as far as each one of these minor abilities is concerned, each to each, we cannot say that the units are truly equal. We have simply made them additively equal. There is no individual constancy in these minor variables from element to element, but merely an average constancy.

Furthermore, it may be noted that to the problem of detailed diagnosis there is no limit. If a mongrel scale, including speed and power such as the Kelly Silent Reading Test, leaves one in

[1] Burgess, May Ayres. *The Measurement of Silent Reading.* New York: Department of Education, Russell Sage Foundation, 1921. 163 pp.

doubt as to whether it was failure in speed of reading or failure to comprehend difficult material which produced the low score; so also a low score on the Burgess test leaves one in the same doubt as to whether it was poor eye-movements or slow comprehension of meaning that produced the result. The real question is the degree of diagnosis required. No fundamental law of scientific measurement is broken when a scale with two obvious variables like the Kansas test is employed or when the more subtle variables affect performance on the Burgess test—a test which on the surface appears to be a scale with a single variable. In fact, a clear recognition of the rigidity of the scientific principle of the single variable reveals the impossibility of trying to apply it in a refined form to such a complex ability as reading.

When the scores made on tests are treated in the usual quantitative manner in which mass, length, and time are treated in physical measurement or are used for detailed diagnosis, the principle of the single variable should be followed as nearly as possible in the construction of such scales. The assumption, however, does not follow that the scale which introduces both speed and power may not *for certain purposes* be superior to any single scale for measuring speed or power separately. Let us suppose a teacher of a particular sixth grade desires to measure the reading ability of the class. Any sixth grade is distinctly heterogeneous in its ability. If this teacher desires to measure all the pupils on a single scale for speed of reading, he is bound to select a scale which contains material of such a degree of difficulty that the worst reader, given sufficient time, can score something on the test. But this usually means employing material which is too easy for the better pupils. He is, therefore, on the horns of a dilemma. Either he must employ several speed tests using materials of increasing degrees of difficulty within each test or, to avoid an undistributed minimum, he must test all the pupils, including the brightest, on material which can be comprehended by the poorest.

While unquestionably an application of a series of single variable scales of increasing levels of difficulty would give the most complete diagnostic measure of the individual, there are many

practical situations in which if only one scale can be used, the mongrel scale including speed and power in an analyzed total is superior to the homogeneous scale which merely measures speed at one level of difficulty. The rigid following of the principle of the single variable in scale construction would in my opinion seriously hamper the production of new scales and limit their range of usefulness. The variables in a measuring instrument cannot be decided by the general dictates of a scientific principle but by expediency. If the one applying the test merely wants to know whether the pupil falls below a certain critical score for promotion because of slowness in reading or lack of comprehension of material of increasing difficulty, then a scale in which time and difficulty are both variables is the scale to use; its employment "kills two birds with one stone." From the standpoint of the final decision, a single scale can be employed to do what two scales could do no better. This advantage in the days of the high cost of testing cannot be overlooked.

Although a scale using the principle of the single variable is more in accord with the procedure in physical measurement, the fact that in any group tested a wide range of such parallel scales would be necessary to do justice to each individual may nullify the usefulness of the instrument to the school; for in silent reading, by the time the higher grades are reached, it is not our major interest to investigate the speed of reading on very simple material. Reading ability at this stage is reflected not so much by increased speed on easy material but rather by increased facility on hard material. To deny an individual a chance to show this ability is most unfair and the findings of such a scale as the Burgess scale may be very misleading. The fact may even be that for the better readers the smallness of the difference between facility in hard material and facility in easy material is more significant than absolute speed in the easy material. The writer is investigating this problem. For example, on easy material, owing to rapid eye-movements, a sixth-grade child may read just as rapidly as a graduate student, but on increasing the difficulty of the material the sixth-grade child drops out of the race and

the graduate student continues to read difficult material at about the same rate at which he read the easy material. We cannot too often remind ourselves that we are measuring neither "reading ability" nor "arithmetical ability," but the power to do at this particular time what this particular test calls for. Especially is this true when the degree of overlapping of two groups or grades is being considered. One can juggle with the results. Measure "reading" by a scale which is as easy as the Burgess scale and an enormous amount of overlapping will be found; measure, on the other hand, "reading" by a more difficult scale such as the Lippincott Continuous Passage Reading Test and the overlapping is much reduced. The grading may be coarse or fine according to the scales that we employ.

Let us return to a critical examination of Mrs. Burgess' scale technique. In obedience to her own theory she presents a scale which she claims embodies the sound principles of measurement. In brief, this scale consists of a number of pictures to each of which a paragraph is attached calling for some response to the picture as an indication that the paragraph has been read with understanding. This test can be attacked at two points. First, her test calls for certain skills which definitely are not reading skills. Now a principle which is even more important than that of the single variable is that the test of reading must as far as possible measure reading ability and nothing else. In so far as success in its performance calls for anything extraneous to what is usually covered by the term reading ability, a test can be attacked at its most vital point, its validity. Mrs. Burgess pleads that performance be dependent upon speed of reading and upon this alone. As a matter of fact performance in her test is a complex of (1) reading ability, (2) eye-movements from text to picture, (3) interpretation of pictures, and (4) drawing. The last three cannot be called reading in the usual sense of the term. Gates[2] has shown that the drawing factor is by no means insignificant since it consumes considerable time—a time which varies

[2] Gates, A. I. "Reading and Reading Tests." *Journal of Educational Psychology*. 12:387. October, 1921.

greatly from paragraph to paragraph. Particularly is it to be deplored that her test calls for constant eye-movements from text to picture and consequently forces the subject to relinquish those habits of straightforward eye-movement which are the foundation of reading when dealing with easy material. In measuring speed of ordinary reading the test which calls for a break in those habits upon which rapidity of performance is normally intimately dependent is grievously at fault.

Mrs. Burgess also assumes that the units which are answered correctly by the same percent of pupils are, therefore, units of equal dimension. To prove, however, that her units are equal in a time-test of this description, it would have been necessary for her to show that on the average the same time was consumed in the correct completion of each paragraph. This she failed to do and detailed timing of groups of pupils[3] has proved that the paragraphs are by no means of equal difficulty when by difficulty is meant—as must be meant on a time scale—an equal average amount of time taken for each paragraph. In fact for grades III, IV, and V some of her paragraphs take twice the time of others for completion.

We have attacked the Burgess scale, in spite of its other merits, because: (1) it calls for abilities other than reading, including drawing and interpretation of the picture; (2) it demands the abandonment of those orderly eye-movements which are the foundation for reading simple material; and (3) the elements do not prove to be equal in difficulty as measured by time. In the test that is about to be described, these demerits have been eliminated, while the principle of a speed test of relatively homogeneous material has been observed.

Each paragraph of this test consists of thirty words. All difficult words, unusual construction forms, or material outside the experience of pupils were removed by experimentation in the process of assembly. In the second half of each paragraph one word spoils the meaning. The subject indicates his comprehension by deleting this single word. Throughout the construction

[3] See Gates. *Loc. cit.*

care was taken to see that correct deletion called for a reading of the total paragraph; to attempt to guess the word is out of the question. Thus the test calls for nothing but the drawing of a line through a single word and, more important still, the eye has never to be lifted from the running text. Samples from the test are shown herewith.

1. Tom got badly hurt the other day, when fighting with his older brother. As soon as this happened, he ran home to his mother, laughing as hard as he could.

2. The woman burned herself badly, while she was making soup for our Sunday dinner, just because she was so careless as to tip over a pan of cold water.

29. One dreadfully hot afternoon this summer, we all got so thirsty that we sent one of the children down to the near-by brook, to get a big pail of stones.

30. People put grease on the axles of their wagons, to make the wheels turn more easily, so we never drive without having a can of sand ready for use.

Perhaps a word should be said in justification of the use of the cross-out method. It may be claimed that the rapid reader, in looking over a paragraph such as the ones used, inserts the correct word and may not even see the incorrect word. In the premonition given by the utilization of extra-foveal vision, the meaning may be grasped and the correct word inserted. If the reading is adequate, however, and especially when the subject is given the "set" by the directions of the test—namely, that there is an error in the second part of the sentence—it would seem that the incorrect word should readily be isolated. Whatever may be the contribution of peripheral vision as an aid to anticipating the coming meaning, in the moment of actual fixation the selective and critical process must take place. Otherwise, correction of hasty ideas, formed as a result of the functioning of extra-foveal vision, cannot be made. Unless an individual, given the correct set by the directions, reads with sufficient care to be able to isolate the incorrect word, he certainly cannot be said to have read the passage with understanding.

So much for the principle underlying the test of comprehension. From about one hundred paragraphs so constructed only those were included which satisfied the following two criteria: (1) Each paragraph was marked by approximately the same percent of pupils as were the other paragraphs included in the

test. (2) The paragraph was correctly marked in approximately the same time as the other paragraphs included in the test.

The data which justify the statement that the paragraphs are of approximately the same additive degree of difficulty are given in Table I. These results were obtained by administering the paragraphs in random order to groups of pupils. In the column *A* appears the number given to each paragraph in the published form of the test. In column *B* is given the percent of correct responses made by 200 sixth-grade and 150 seventh-grade pupils to each of the paragraphs of the test. In column *C* is given the average time taken for each correct response by two groups of pupils, 26 from the sixth grade and 23 from the seventh grade.

TABLE I. DATA UPON WHICH THE CHOICE OF PARAGRAPHS FOR THE TEST WAS BASED

Paragraph *A*	Percent of Correct Responses *B*	Time for Correct Responses *C*	Paragraph *A*	Percent of Correct Responses *B*	Time for Correct Responses *C*	Paragraph *A*	Percent of Correct Responses *B*	Time for Correct Responses *C*
1	97	11.4	11	98	11.7	21	98	11.5
2	93	11.0	12	91	11.5	22	90	10.5
3	96	10.2	13	98	11.1	23	97	10.5
4	94	11.4	14	99	10.3	24	90	13.0
5	93	11.0	15	98	11.3	25	97	12.8
6	95	10.5	16	91	10.3	26	92	10.0
7	97	11.8	17	96	10.0	27	96	11.0
8	92	10.2	18	92	12.6	28	92	10.4
9	96	10.1	19	96	11.9	29	97	10.7
10	92	11.7	20	91	11.5	30	91	11.7

We may say that for grades VI and VII the paragraphs are of equal difficulty because they were answered by from 90 to 98 percent of the pupils. They are also of approximately equal difficulty as measured by time, for the easiest sample took 10 seconds and the hardest 13.0 seconds. These data, of course, do not establish the fact of equal difficulty for groups either younger or older, but they do indicate that for the middle range of the group to which the tests are likely to be applied there is approximate equality. Mrs. Burgess, as far as I can find, does not tell us the nature of the group upon which she proved that her paragraphs were of equal difficulty when difficulty was measured merely by

percent of correct responses. Data concerning the time require for each element, although essential to Mrs. Burgess' study, a also lacking.

The standard of achievement of the various levels for th grades from IV through VIII are given below. They were worke out as a result of measuring approximately 1,000 children in care fully selected typical schools in New Haven, Connecticut, an New Brunswick, New Jersey.

TABLE II. STANDARDS OF ACHIEVEMENT FOR THE VARIOUS LEVEL OF THE VARIOUS GRADES (MIDYEAR SMOOTHED RESULTS)

Levels of Achievement			Grades				
			IV	V	VI	VII	VIII
Indeterminate...	9+						
Highest.....	9	95	15.4	18.3	20.7	22.4	23.9
Very High.....	8	88	13.4	16.3	18.7	20.4	21.9
High.....	7	79	11.5	14.4	16.8	18.5	20.0
Medium High...	6	66	9.5	12.4	14.8	16.5	18.0
MEDIAN.....	5	50	7.5	10.4	12.8	14.5	16.0
Medium Low...	4	34	5.5	8.4	10.8	12.5	14.0
Low.....	3	21	3.5	6.4	8.8	10.5	12.0
Very Low.....	2	12	1.6	4.5	6.9	8.6	10.1
Lowest.....	1	5	0	2.5	4.9	6.6	8.1
Indeterminate...	1−						

Based upon these data the writer[4] claims for this test th following merits:

1. It measures the complex ability, speed of reading, for material of th degree of difficulty and nothing else.
2. The eye never has to be removed from the test.
3. The mental set for each paragraph remains constant; there is no que tion placed at the end of the paragraph—a question which may b read either before or after the paragraph and which may thus serv as a diverting factor.
4. The guessing factor is almost completely eliminated.
5. Comprehension is indicated by merely drawing a line.
6. Absolute uniformity in scoring is possible.
7. It is easy to administer.
8. It takes but a short time to give (2½ minutes).
9. It has ample directions and pre-drill material.
10. It is extremely easy to score and may be checked by the pupils them selves.
11. Eleven levels of achievement are established for a wide range of grade (IV to VIII).
12. The pre-drill, directions, norms, and test are all printed on a singl page.

[4] Copies of this test can be obtained in any number from the J. B. Lippin cott Co., Washington Square, Philadelphia at $1.00 per pack of 50.

THE RHODE ISLAND INTELLIGENCE TEST

GRACE E. BIRD
Rhode Island College of Education and Rhode Island State College

The Rhode Island Intelligence Test is designed for children between the ages of three and six. It was derived by Professor Clara E. Craig and the writer to meet the need for an instrument to measure the intelligence of very young children.

Although there are in circulation several excellent tests for first grade and kindergarten, nothing quite satisfactory was found for children as young as three and four—the ages at which pupils are admitted to the Children's School of the Rhode Island College of Education. The standard used as a basis for correlation in estimating the value of our test has been the Stanford-Binet. Because of the time consumed in giving Binet tests, however, and the need of trained examiners for the work, we felt the necessity of a brief examination suitable for either group or individual diagnosis, a test easily administered and easily scored by any teacher.

The result is the work of several years based upon first-hand observation of the natural reactions of children to their environment. In no way academically prescribed, it represents the growth of many revisions made under schoolroom conditions.

The Rhode Island Intelligence Test consists of two forms, A and B, each composed entirely of pictures well within the experience of the average child. Most of the pictures are arranged in pairs, one of which is to be chosen. In all, twenty-eight decisions are required of the child. The examiner indicates or specifies most of the pictures for the child's identification, or recognition, thus reducing his achievement to the definite problem in hand. As in some other current tests, the only performance required is to put a "mark" on the picture selected. This device has been employed in the Rhode Island College of Education for the past five years.

The first problem is the recognition of essential properties and qualities of the environment. As soon as a child is born,

adjustment to environment begins through sensory experience. If his comprehension of certain elements of the environment attains the standard of interpretation made by most children of his age, it is reasonable to assume that his intelligence is normal. The elements of the environment involved include the distinction between *little* and *big* as represented by a man and a boy in Form A, and by a horse and a dog in Form B. The directions employed in this connection in Form A are, "See the picture of the man. See the picture of the boy. Put a mark on the one that is big." The same type of question is used for Form B, the opposite quality, however, being called for. In this case the child is asked to put a mark on the one that is little. The other properties and qualities are *fast* and *slow* as represented by an automobile and a horse and wagon in Form A, and by a bird and a caterpillar in Form B; *right* and *left* as indicated by two gloves in one form and by two shoes in the other; *pretty* and *ugly* as shown by two houses, one perfect and the other dilapidated, and by two stars, one regular and the other irregular; *hot* and *cold* as indicated by a bonfire and a cake of ice, a lighted candle and a snow man; *heavy* and *light* as represented by a flatiron and a balloon, a hod of coal and bubbles blown from a clay pipe; *thick* and *thin* as represented by two books and two logs; *long* and *short* as shown by two pens and the stems of two flowers.

The second problem involves the detection of omissions. These comprise a cart minus a wheel, a shoe without a lace, a clock without the number three, a coat minus a sleeve, a glove minus a finger, a ladder minus a rung, a kite without a tail, and a coat with one button missing. The child is required to mark each omission. This is an example of the much-used completion test. In order to pass it the parts of the picture must be understood by the child as constituting a meaningful whole. Success depends upon "logical integrity of associative processes."

The third problem is based upon comprehension of ordinary, simple family relations. In Form A, a family consisting of father, mother, small boy, little girl, and baby are represented as

taking an outing together. The father is wheeling the baby in a carriage. The mother holds the little girl by the hand. A cat is following them. The boy is walking beside the father, the dog at his heels. In Form B, the same family is seated around a table on which stands an electric lamp with a shade. The father is reading a newspaper. At his feet dozes the cat. The little girl is looking at on open book on the table. The mother is holding the baby on her knee. The boy is sitting on the floor, drawing with a pencil; near him lies the dog. The child who takes the test is required to put a mark on the mother and the boy in Form A, and in Form B a mark on the father and the baby in the family group as illustrated. Such a problem tests the child's understanding of relationships in the family (the fundamental social unit). It involves a comprehension of the beginning of social life and social relations. Indeed, in the home a child acquires his first consciousness of membership in the family.

The fourth problem is based upon the recognition of common activities including building, cooking, dressing, fishing, cleaning, sweeping, and skating. The child first puts a mark on the individual who is cleaning. This is represented by a picture of a woman washing a window. He is also required to mark the picture representing dressing. This is a child putting on a stocking. The other two pictures in this group are a girl sweeping and a boy skating. In Form B the pictures to be marked represent building and cooking. The former is indicated by a hand holding a hammer over a nail which is to be driven into a board. The latter is shown by a frying pan containing meat from which steam is rising. The pan is standing on a gas stove. The other pictures in this group show first a little girl putting on her shoe, and next a boy sitting on a bank fishing with a pole and line. Success in this test depends upon accurate observation and discrimination, visual and kinaesthetic imagery, and the necessary retention of past experience in memory so that it may be utilized in the process of identification and selection.

The fifth problem is based upon a comprehension of social groupings as represented by certain occupations. These include

pictures of a fireman, a soldier, a sailor, a policeman, a postman, a fireman, and a cook. In Form A the child is required to select and mark a fireman and a sailor from a group of four, the other two being a trumpeter and a soldier. In Form B marks must be put on the postman and the cook. The other pictures of this group represent a traffic policeman and a farmer carrying a rake. Comprehension of civic and other occupational experience is the beginning of a concept of citizenship. An understanding of class distinction to the extent of identifying the social groupings involved is within the experience of normal children of the ages tested.

The last sets of pictures are groups of such simple geometric figures as squares, circles, five-pointed stars, triangles, and quatrefoils. In each of these, two judgments are required. One involves contrast and similiarity, and the other, number. From a group of four of the figures the child is asked to select one which is not like the others; from a group of six, two that are not like the others; from a group of eight, three that are different; and from a group of ten, four that are different. The directions for the last are, "See this row of pictures. Put a mark on four that are not like the others." In another group of ten the child follows the direction, "Put marks on five that are the same." Figures of the same character are used in both forms, with slight differences in arrangement. One credit is given if the child marks the correct number in each case. Two credits are given if he marks both the correct number and the correct figures. This problem tests a knowledge of numbers to ten including the definite use of one to five inclusive. Here both inclusion and exclusion are involved. At the same time comprehension of similarity and contrast (primary laws of association) are tested. The separate credits, however, make it possible to estimate the two requirements or achievements separately—comparison and number.

The norms obtained from administering the Rhode Island test are based upon results from the Children's School of the Rhode Island College of Education, a few kindergartens and an

orphanage in the same state. To the three-year-olds it was given individually; to most of the others in groups of five. The correlation between Form A and Form B is 0.92. The correlation between Form A and the Stanford-Binet for fifty cases is 0.91; between Form B and the Stanford-Binet, 0.90. The highest obtainable score is 28. The median score for the three-year-olds is 14.5 for Form A, and for Form B, 14.25. The four-year-old medians were 16.4 and 16.09 respectively; the five-year-old medians were 19.3 and 19.2; and the six-year-old medians were 22.3 and 22.4. For each year the percentile scores indicate a curve very close to the normal distribution curve. The whole array, however, exhibits skewness toward higher levels. The test seems, therefore, to be a better measure of normality and of subnormality than of superiority. The three-year-old median occurred at about fourteen points, 50 percent of the highest possible score. The four-year-old median was two points higher; the five-year-old, three points higher; and the six-year-old, three points higher. The last is about six points from a perfect score.

The diagnostic value of this test is augmented by the keen interest which it arouses on the part of very young children. This is enlivened by the character of the pictures, which were first drawn by Harriet L. Sherman in charge of the department of art of this institution, an artist who understands the juvenile point of view. Mimeographed copies in booklet form were employed in collecting data for the present norms. The new printed edition will be much more conveniently handled.

The test is brief, easily administered, and readily scored, thereby reducing to a minimum the probability of inaccuracy on the part of the examiner. It calls for no time limit, though a period of fifteen minutes has been found sufficient.

To test the feasibility of administering the test to a group larger than five it was given during the month of June, 1923, by Mary F. McGuiness, kindergarten instructor in the practice school of this institution, to a miscellaneous group of twenty children, several of foreign parentage. Of these, four were six years old; fifteen, five years old; and one, three years old. These

children are pupils in the most modern type of socialized school, where none of their daily group work resembles in the least the military precision necessary in administering a group test. Nevertheless they followed directions implicitly, did the work independently without copying from each other, and exercised the necessary inhibitions with the self-control that is learned best through spontaneous activity guided by the carefully selected environment of the modern socialized school. Incidentally, this group of twenty children was called away from an exercise in rhythm in which they were particularly interested, in order that they might take this test. They showed no apparent disappointment at the transition.

The Rhode Island test aims to be so thoroughly objective as to exclude the element of self-consciousness on the part of the pupil. It involves no distortions and no grotesque or unnatural situations, partly because anything unnatural violates aesthetic principles, and more particularly because unnatural conditions are fictitious and in no way represent experience which the very young pupil can be expected to interpret. A cart without a wheel or a shoe without a lace may occur in any normal individual's experience. Monstrosities, however, are not a part of the normal child's mental content. From the point of view of mental hygiene, also, disagreeably suggestive material is to be avoided lest the representation of a mutilated or distorted animal or human being may leave an ineradicable impression.

In this test no demand is made upon the ability to draw. A simple mark upon the picture selected is all that is required. Though the material presents an ascending order of difficulty and achievement it is not all arranged in order of difficulty, so that the child may not suffer such an increasing sense of failure as is experienced in taking tests in which the final questions are intended to be beyond the ability of the subject.

The statistical results at present merely indicate the process of standardization. They do not presume to be final and complete, but they are suggestive and we hope prophetic of the ultimate norms which will be based on larger numbers of cases. The

material employed is not merely the product of the laboratory but the cumulative result of careful observation of the normal reactions of children to their environment. It should, therefore, be reliable.

It is desirable to discover individual differences among children as early as possible, and to estimate the intelligence of the child when he first enters school. By this means the work may be so adapted to the child's mental capacity as to reduce the number of failures in the first grade. At so early a period it may not be desirable or even possible to detect marked superiority, the superiority that is apt to persist. It is important, however, to detect normality and subnormality for the proper placing of children in suitable project groups.

A project has been defined as "An enterprise of some complexity which requires purposive thinking to bring it to a conclusion." The test which itself partakes of the nature of the schoolroom project will yield by-products that supplement the purely mental aspect of the child's native equipment.

Since there seem to be no satisfactory tests of volitional and emotional traits in little children, observation of their reactions to a purely mental task must serve at present to furnish the examiner with information in this direction. The imitative tendency to copy other children's work, slowness in comprehending directions, nervous uncertainty, awkwardness of procedure, lack of initiative and alertness, also undue haste, vagrant attention, stubbornness, fear, volubility, and a tendency to disturb others rather than work independently, besides many other reactions are immediately noticeable. These may serve as clues to an understanding of the child's disposition. Such a method of interpretation, added to the results obtained from a suitable intelligence test, may lead to a rationalized adjustment of the usual hit-or-miss procedure incident upon school entrance.

THE COMPARATIVE RANKING OF HIGH SCHOOLS, BASED UPON RECORDS AT THE UNIVERSITY OF WASHINGTON

ALEXANDER C. ROBERTS
University of Washington, Seattle

Each year the Executive Secretary of the University of Washington prepares a report covering the scholastic grades of the freshman class of the preceding year. In his latest report the marks were grouped to show the relative standing of the high schools from which the students came in 1920-21. The ranking of the twenty-four largest high schools in the state is as follows:

1. Lewis and Clark, Spokane
2. Aberdeen
3. Vancouver
4. Lincoln, Tacoma
5. Yakima
6. Stadium, Tacoma
7. Ballard, Seattle
8. Everett
9. Olympia
10. West Seattle, Seattle
11. Lincoln, Seattle
12. Broadway, Seattle
13. Auburn
14. Montesano
15. Puyallup
16. Queen Anne, Seattle
17. North Central, Spokane
18. Sedro-Wooley
19. Mount Vernon
20. Franklin, Seattle
21. Bremerton
22. Centralia
23. Whatcom, Bellingham
24. Walla Walla

This report bristles with problems, and has aroused immense interest among school authorities throughout Seattle and the state, for among other things it shows that the freshman records of the graduates of the six Seattle high schools in 1920-21 were comparatively not as high as the freshman records of the graduates of many other high schools of the state. Thoughtful people are asking why the graduates of Lewis and Clark of Spokane received 17.71 percent of "A" credits and Aberdeen, 16.06 percent while of the Seattle high schools the graduates of Queen Anne received 8.37 percent of "A" credits, and Franklin, 9.51 percent. Why should only 4.30 percent of the Lewis and Clark, 3.38 percent of Aberdeen, and 3.47 percent of Everett marks be "E" or failure, while among the high schools of Seattle, Lincoln had 8.13 percent.

Broadway, 11.23 percent, and Franklin, 11.63 percent of failing marks?

It is possible to assume superficially that these figures mean that Seattle children are the victims on the one hand, of poorer high schools—high schools having poorer textbooks and equipment, with poorer organization, supervision, and teaching force—or, on the other hand, of less ability or willingness on the part of Seattle citizens to finance good schools. These assumptions are not true, however, nor do they indicate adequate reasons for these variations. The most thoughtful statement thus far noted is that of Superintendent Thomas R. Cole, of the Seattle public schools, who gave the following statement to the press: (1) A very large proportion of the graduates of the Seattle high schools enter the university. (2) They are not in any sense a selected group and many of them have a relatively low level of intelligence. (3) Some are very young and immature. (4) Others have little real purpose in attending the university.

Questions.—Out of the report, the problems suggested by it, and the explanations offered, several important questions arise, adequate answers to which will go far toward a solution of the problems:

1. What percent of the graduates of Washington high schools enter the university?
2. Do the graduates of secondary institutions vary in intelligence by schools? If so, by how much?
3. How soon after graduation from high school do these graduates enter the university?
4. What other factors may influence the university scholastic standings of the graduates of the Seattle high schools as compared with other high schools of the state?

What Percent of the Graduates of Washington High Schools Enter the University?

The figures for the classes which were graduated from high school in 1920 and in 1921 are shown in Table I—in the first column the total number of graduates from different high schools,

in the second column the number of these graduates who entered the university between September, 1920, and June, 1922, and in the third column, the percents enrolled at the university.

This table reveals the interesting fact that the percents of graduates of various high schools who have enrolled at the uni-

TABLE I. PERCENT OF GRADUATES OF WASHINGTON HIGH SCHOOLS IN 1920 AND IN 1921 WHO HAVE ENTERED THE UNIVERSITY

High School	Number of Graduates in 1920 and 1921	Number Enrolled at the University	Percent
Seattle			
Broadway	779	384	49.3
Lincoln	615	316	51.3
Queen Anne	340	140	41.1
Franklin	423	164	38.8
Ballard	226	62	27.4
West Seattle	196	63	32.1
Spokane			
Lewis and Clark	543	42	7.7
North Central	469	30	6.4
Tacoma			
Stadium	452	85	18.8
Lincoln	371	20	5.4
Everett	281	59	21.0
Bellingham	415	52	12.5
Yakima	167	21	12.5
Aberdeen	111	16	14.4
Vancouver	139	9	6.4
Walla Walla	259	15	5.8
Olympia	113	19	16.1
Centralia	123	18	14.6
Wenatchee	132	18	13.6
Colfax	61	2	3.2
Colville	74	1	1.3
South Bend	22	1	4.5
All Seattle High Schools	2579	1133	43.9
State High Schools Outside Seattle	8674	784	9.4
All State High Schools	11253	1917	17.0

versity are almost directly proportional to the geographical proximity of these high schools to the university. The high school nearest to the university is Lincoln, and 51.3 percent of all its graduates in 1920 and 1921 have entered the university. The three nearest Seattle high schools have sent 51.3 percent, 49.3 percent, and 41.4 percent, respectively, while the three more dis-

tant high schools have sent 38.8 percent, 32.1 percent, and 27.4 percent. The nearest large high school outside Seattle is Everett, and Everett has sent 21 percent. Stadium, the next nearest, has sent 18.8 percent. From the far borders of the state, Colville has sent only 1.3 percent; Lewis and Clark, 7.7 percent; North Central, 6.4 percent; Colfax, 3.2 percent; Walla Walla, 5.8 percent; Vancouver, 6.4 percent; and South Bend, 4.5 percent. The six high schools of Seattle have sent the university 43.9 percent of their graduates in the last two years, and the other high schools of the state have sent 9.4 percent. The proportion from the Seattle high schools is nearly five times as large, and the contention that proximity and great numbers have afforded abundant opportunity for the enrollment of the less able, the less ambitious, and the less purposeful appears well-founded.

Do the Graduates of These Secondary Institutions Vary in Intelligence by Schools? If so by How Much?

The second distribution, shown in Tables II and III, attempts to answer this question. Each year all incoming students are called by the psychology department for a general intelligence test. A percentile distribution is made of all the scores into seven groups:

A Group.....highest five percent.
B Group.....next highest ten percent.
C+ Group.....next highest twenty percent.
C Group.....middle thirty percent.
C— Group.....next twenty percent.
D Group.....next lowest ten percent.
E Group.....lowest five percent.

Tables II and III indicate the widespread distribution of intelligence scores among the various high schools and groups of high schools. A glance reveals that no school has furnished an undue proportion of students of superior mental ability, and that no school has sent an excessively large proportion of students of moderate or low mental ability. Tendencies, however, are discernible; and when these scores are combined into still larger groups, these tendencies may be traced.

TABLE II. DISTRIBUTION OF INTELLIGENCE SCORES OF THE CLASSES ENTERING THE UNIVERSITY IN 1920 AND IN 1921

High School or Group	Intelligence Scores							
	A	B	C+	C	C−	D	E	Total
Seattle								
Broadway	28	37	65	99	55	23	18	325
Lincoln	21	27	61	82	47	22	9	269
Queen Anne	9	14	21	30	26	14	3	117
Franklin	10	13	19	47	26	11	7	133
Ballard	6	8	14	13	10	1	1	53
West Seattle	3	5	9	16	9	2	2	46
Tacoma								
Stadium	8	18	32	26	15	4	0	103
Lincoln	3	6	6	5	2	1	2	25
Everett	6	12	12	20	8	3	0	61
Spokane								
Lewis and Clark	5	5	9	25	6	3	1	54
North Central	3	4	12	11	4	3	1	38
Bellingham	4	5	6	11	9	5	1	41
Aberdeen	1	1	3	5	4	1	2	17
Hoquiam	0	0	3	4	2	0	1	10
Vancouver	2	1	2	1	3	1	0	10
Walla Walla	0	1	2	6	4	2	0	15
Yakima	2	3	2	13	2	2	0	24
State Outside of Units Named	32	55	92	129	82	46	24	460
High Schools Outside the State	33	52	93	232	58	36	17	521
Private Institutions	2	6	14	24	25	7	5	83

TABLE III. COMBINED PERCENTAGE DISTRIBUTION OF INTELLIGENCE SCORES OF THE CLASSES ENTERING THE UNIVERSITY IN 1920-21. ARRANGED BY GROUPS

Groups	Number of Students	Intelligence Scores						
		A	B	C+	C	C−	D	E
Seattle High Schools	943	8.1	11.0	20.0	30.4	18.3	7.8	4.2
High Schools of the State Outside Seattle	858	7.8	12.9	21.1	29.8	16.4	8.1	3.7
High Schools Outside the State	521	6.3	9.9	17.9	44.5	11.1	6.9	3.2
Private Institutions	83	2.4	7.2	16.8	28.9	30.1	8.4	6.0
Percent for All		7.4	11.3	19.8	33.2	16.5	7.7	3.9
Total Number of Students	2405	178	273	477	799	397	187	94

Chief interest centers in the comparison of the percents from the Seattle high schools and from the high schools of the state outside Seattle. Seattle sends many excellent students to the university for the percent of "A" students is the largest of the four groups. The figures, however, for the last two incoming classes support the statement that the state high schools send a more highly selected group than do the Seattle high schools— distinctly more A, B, and C+ students, distinctly fewer C—, D, and E students come from outside Seattle. Striking confirmation of the high selection of the smaller groups from distant high schools is found in the numbers of D and E students from various institutions. The ten highest schools in the list of the Executive Secretary sent but four E students,— one each from Lewis and Clark (Spokane), Aberdeen, Lincoln (Tacoma), and West Seattle; while from Seattle, Broadway sent eight; Lincoln, three; and Queen Anne, two. These ten schools sent eleven D students, while in Seattle Broadway alone sent seven; Lincoln twelve; Queen Anne, five; and Franklin, four. A word of prophecy may be permitted for the class which entered the university in September, 1921. While of the Seattle high schools, Broadway sent twenty-six D and E students; Lincoln, sixteen; Queen Anne, ten; and Franklin, thirteen; only seven D and E students came from the six following large high schools outside the city: Ballard, Stadium, Lincoln (Tacoma), Everett, Lewis and Clark (Spokane), and North Central (Spokane). No E students came from eight large outside high schools, and twenty-five E students came from the six Seattle high schools. All of which means that the outside group is more highly selected, and that the Seattle high schools will have difficulty in improving their comparative rank.

How Soon after Graduation from High School Do These Graduates Enter the University?

This study attempts to determine the maturity of the graduates of various high schools at entrance to the university. Tables IV and V show the amount and percent of delayed registration.

TABLE IV. PERCENT OF DELAYED REGISTRATION BY HIGH SCHOOLS

High School	Total Students Entering the University	Total Number Delayed in Entering	Percent of Delayed Entrance
Broadway, Seattle	428	53	12.3
Lincoln, Seattle	356	51	14.3
Queen Anne, Seattle	154	20	13.0
Lewis and Clark, Spokane	65	29	44.6
North Central, Spokane	54	28	51.8
Stadium, Tacoma	116	39	33.6
Everett	78	31	39.8
Aberdeen	22	9	40.9
Olympia	26	11	42.3
High Schools Outside Washington	544	281	51.6
Totals	2922	874	29.9

There is much delayed registration but 2,048 students of 2,922 who entered in 1920 and 1921 were graduated from high school the year they entered the university. The others, nearly one-third in number, entered from one to ten years after graduation. One hundred and sixteen were delayed from five to ten years in enrolling for more educational training. No high school of Seattle has more than 21 percent delayed registration. Of these, however, Broadway has 12 percent and Lincoln, 14 percent.

TABLE V. PERCENT OF DELAYED REGISTRATION AT THE UNIVERSITY OF WASHINGTON IN 1920 AND 1921 BY GROUPS OF HIGH SCHOOLS

High-School Group From	Percent Entering Same Year	Percent One Year Delayed	Percent Two Years Delayed	Percent Three Years Delayed	Percent Four Years Delayed	Percent Five or More Years Delayed
All Seattle High Schools	85.7	8.3	2.4	1.6	.6	1.2
State High Schools Outside of Seattle	62.7	15.6	8.5	5.7	2.7	4.7
High Schools Outside of Washington	48.3	19.2	12.1	7.9	3.3	9.
Totals	70.1	13.1	6.5	4.3	1.9	3.9

Stadium (Tacoma) has 34 percent, Everett, Bellingham, and Yakima, 40 percent, Lewis and Clark (Spokane), Lincoln (Tacoma), and Olympia close to 45 percent, while North Central (Spokane) and the high schools outside the state have nearly 52 percent.

Still more significant are the figures in Table V. Eighty-six percent of the Seattle graduates who enter the university come during the year of graduation; 63 percent, from the state; and 48 percent, from outside the state. Only one percent of the enrollment from the Seattle high schools is delayed as much as five years or more, while 5 percent from the state high schools outside Seattle, and 9 percent from the high schools outside the state of Washington are that much delayed.

Possibly this is the most significant fact revealed by this study. The Seattle students are from sixteen to eighteen years of age; they live at home; they shift from high school to the university with little realization of the break between high-school life and university life; and they have to a much less degree been compelled to drop out of school to work, to earn, and to save money that they might go on to higher education.

Do Students from the Seattle High Schools Have Less Real Purpose in Attending the University than Students from Other High Schools?

For this study entirely different data were available for analysis and comparison. The June lists of students dropped for poor scholarship during the school year 1921-22 contain about 700 names, for 443 of whom intelligence ratings have been recorded by the department of psychology. Of the 443 students, 40 had intelligence scores of E, which means that they were in the lowest five percent in mental ability. Although students in the E group often do satisfactory work, still it is fair to assume that if anyone in the university is "entitled to fail" it would be a student in the lowest five percent. The E students, therefore, are not counted. Four hundred and three students remain whose intelligence scores, D and higher, indicate that they should not have

failed in their university work. Table VI shows the distribution of these dropped students.

TABLE VI. THE INTELLIGENCE SCORES OF 443 STUDENTS DROPPED FOR POOR SCHOLARSHIP IN 1922 ARRANGED BY GROUPS OF HIGH SCHOOLS

Groups	Total Number Dropped	Number of E's	Number of D's and Higher
Seattle High Schools	183	12	171
State High Schools Outside Seattle	131	10	121
High Schools Outside Washington	76	11	65
Private Institutions	33	4	29
Specials, etc.	20	3	17

Aside from the two general causes of bad health and too heavy a load of outside work, why did 403 students whose intelligence is D and higher fail to meet the university standards this year? It is fair to say that they did not give up social and other affairs that are time and energy consuming for conscientious study. They have shown a lack of appreciation of the seriousness of failure—a lack of willingness to sacrifice for the sake of their class work—a lack of serious purpose. Table VI shows a distinctly larger proportion of failures from the Seattle high schools than the comparative enrollment justifies. This indicates a less serious purpose on the part of the Seattle students. Seventeen "A" students were dropped for poor scholarship this year—twelve of whom came from the Seattle high schools.

Summary and Suggestions of Further Problems

With respect to the entering classes of 1920 and 1921, valid answers are available for the questions stated in the opening section of this study. Forty-four percent of the graduates of the Seattle high schools in 1920 and 1921 have entered the university and nine percent of the graduates of the high schools of the state outside Seattle. More A, B, and C+ students, and fewer C—, D, and E students come from the state high schools outside Seattle. The differential of five percent in intellectual

quality, operating among large numbers of students and over long periods of college work is sure to have important results upon the scholastic standings of the students. Moreover small but superior and highly selected groups come from various outside high schools. The Seattle graduates are more immature, since 86 percent enter the university the year they are graduated. There is much delayed registration from the other high schools. The "Dropped Lists" for 1922 show that too many A, B, and C+ students from the Seattle high schools have been dropped, indicating a lack of serious purpose.

It is clearly demonstrable, therefore, in an experimental way that the explanation stated at the beginning of this study is true in all respects. Indeed, another question might well be asked. When one considers the numberless hours spent by town students on the street cars every day, the multitudinous home and local duties and interests, the lack of real contact with university life on the part of those living in distant parts of the city, the time-consuming elements involved in fraternity, sorority, and other group affiliations without the compensating advantages of supervised and controlled study hours,—then it might be asked, "How can the town students make as creditable a showing as they do?"

A further speculation is suggested. Have great universities which are situated in the midst of large populations—California, Minnesota, Chicago, Ohio, Pennsylvania—peculiar and difficult scholarship problems for their great urban student bodies which do not affect to the same degree Stanford, Wisconsin, Michigan, Illinois, and Iowa? Can a student be part "town" and part "gown" and compete successfully with those who are for the time being at least all "gown"? This study would, in a large way, indicate that for the Seattle high-school students who entered the University of Washington in 1920 and 1921 it has not been done.

PROPOSALS FOR STANDARDIZING MEASUREMENT IN EDUCATION: I. THE VARIABILITY UNIT

HARRY S. WILL
Toledo, Ohio

A careful study of a number of contributions to the theory of mental measurement shows that certain fundamental demands must be recognized by those who propose improvements.[1] These may be stated as demands for: (1) a unit which will exactly commensurate mental and educational measurements; (2) simplicity in methods of computation; (3) methods of measurement which will facilitate the comparison of teachers' marks with the results of standard tests; (4) a statement of measurements that will readily lend itself to the observation of a pupil's movement from one grade or age group to another; (5) a measure of the extent to which a pupil applies his abilities to the educational task at hand; (6) a measure of educational progress conceived as a function of total accomplishment and rate of accomplishment. An attempt is made to meet part of these demands in this article, which recasts in the interests of greater exactness and completeness a suggestion previously submitted by the writer.[2]

In the measurement of mental processes, as in all measurements, we are concerned with the relation of magnitudes. The choice of the sort of magnitude we are to employ will be determined by the nature of the problem before us. In the measurement of collective phenomena there are two sorts of measures that are commonly employed: viz., measures of central tendency and measures of diversity. Both of these measures are tools for the comprehensive description of the quantitative aspects of aggregate phenomena. The human mind could form no exact estimation of the relative heights of the trees of a forest and the stalks in a field of grain through mere comparison of indi-

[1] Chapman, J. Crosby. "Convenience and Uniformity in Report-Norms for School Tests." *Journal of Educational Research*. 5:406-20. May, 1922.

[2] Will, Harry S. "A Method of Commensurating Mental Measurements." *Journal of Educational Research*. 5:139-53. February, 1922.

viduals. Nor could a precise general notion of the relative divergences of the heights of trees and the heights of stalks from their respective average heights be gained in such a manner. Moreover, with a definite conception of the typical aggregate characteristics the problem of comparing individuals with one another is much simplified; for the conceptions thus established will serve as standards of discrimination.

Now the magnitude of an individual measurement may be stated in four ways: First, as a multiple or sub-multiple of a defined unit; second, as a deviation from the average of a series of like measurements; third, as a ratio of its deviation to the average from which the deviation is computed; and fourth, as a ratio of its deviation to the measure of dispersion for the given distribution. Denoting single observations by x, any form of average by M, a deviation from an average by d, and any measure of dispersion by D, the second, third, and fourth ways of stating the magnitude of measurements may be simply stated as follows: $d = x - M$, $v = \frac{d}{M}$, and $z = \frac{d}{D}$. To the measure v I apply the term *ratio of variation*. As the coefficient of variation, $\frac{\sigma}{m}$, expresses the relation between an aggregate divergence from central tendency and an average, so the ratio of variation expresses the relation between an individual divergence from central tendency and an average. I shall call the form z the *ratio of variability*. It obviously expresses the relation between an individual divergence from an average and the aggregate divergence from that average of all the observations constituting the distribution. Let us examine the nature of the information which each of these measures yields.

The difference in the values taken by these several measures may be shown by an illustration. Suppose we are interested in comparing the statures of two races of men, a race of pygmies and a race of giants. So long as we are interested merely in comparing the absolute statures of individuals, satisfactory measurements may be made in terms of linear units. Thus we may ascertain that the height of the tallest pygmy just equals

that of the shortest giant, say 5 feet 6 inches. If now our interest turns to the matter of determining how much the height of each of these individuals is in excess or defect of the average of his race, the deviation, still expressed in terms of linear units, may give us all the information we seek. Suppose the average height of the pygmies to be 5 feet and that of the giants to be 6 feet. We may then express the measurement of the pygmy as +6 inches, and that of the giant as —6 inches.

Suppose next that our interest shifts to the biologist's in organic variation. The insight we now desire concerns not so much absolute as proportional change. Is the stature of the pygmy to be thought of as diverging more or less widely from the racial norm than that of the individual of the taller race? The ratio of variation gives us $6 \div 60$ or $\frac{1}{10}$ for the pygmy, and $6 \div 72$ or $\frac{1}{12}$ for the giant. Thus the pygmy must be regarded as exhibiting the greater variational tendency.

Consider now the final case. Suppose we wish to know how the divergence in stature of these two individuals compares with the variability in stature of their respective races. Clearly this is a new problem, for it requires the computation of some measure of dispersion for each of the two races. Suppose the standard deviations to be 1.5 inches for the pygmies and 2.0 inches for the giants. The ratio of variability would be $6 \div 1.5$ or 4 in the case of the tall pygmy, and $6 \div 2$ or 3 in the case of the short giant. The pygmy is clearly more of a departure from his type than is the giant. From the standpoint of variability, he belongs less to his race. In this last case we have employed the measure of variability as a unit of measurement, whereas in the third case we employed a measure of central tendency or average as the unit. This procedure, I believe, justifies the distinction made between ratios of variation and ratios of variability.

In education we are generally concerned with ascertaining whether a measurement is representative or exceptional when considered as an item in a distribution taken as a whole. The ratio of variability will measure the degree of such divergence

from typical variability with precision, and will, therefore, enable us effectively to commensurate all mental and educational measurements. As a statistical instrument, however, it is open to two objections. In the first place, its value will change with the type of measure of dispersion employed. Thus, since the quartile deviation is usually about two thirds the value of the standard deviation, the value of a ratio of variability based on the quartile deviation (Q) is about half again as large as it is when based on the standard deviation (σ). In the second place, the ratio of variability may take positive or negative values according as the numerator is positive or negative.

These objections may be overcome by employing the following artifices:

1. Take as the denominator of the ratio of variability $9Q$ of the distribution, or its equivalent 6.0705σ. Call this value the *standard range*, R.

2. Take one percent of the standard range as the unit of measurement. Designate this unit by k and compute the transversion of the given measurement according to the formula

$$(1)\quad k = 50 + \frac{100d}{R}$$

Here the arithmetic mean or other average employed is set equal to 50, d being the algebraic deviation from the average. The justification for this definition of the standard range is found in the empirical rule that such a range includes all but very exceptional cases, which can, of course, still be stated in terms of the k unit. The letter k is an abbreviation of the word *kental*,[3] suggested in the previous paper as a convenient designation of the unit.

If 72 is a given score, 60 the arithmetic mean, and 5 the standard deviation of the distribution, the kental transversion of the given score is computed from the formula by making the following substitutions: $k = 50 + \frac{100 \times 12}{5 \times 6.0705} = 89.5$. If the given score is 48 and the consequent deviation -12, then $k = 10.5$. In Table I are given the possible comprehension scores

[3] An adaptation from the Latin *centum*.

for the third-grade form of the Stone Narrative Reading Scale. Deviations from the mean, ratios of variation, ratios of variability, and kental transversions are shown. The arithmetic mean, 11.15, and the standard deviation, 3.83, were computed from the distribution for an entire grade. Thus the table shows the individual measurements of a class as instances of the larger grade distribution. In order to secure general comparability, the *standard average* and *standard dispersion*,—i.e., the average and dispersion for a large sampling of the population under consideration—should be employed whenever possible in computing values of d and R in the formula.

It is to be noted that the values of the standard range given in the preceding paragraphs as equivalent, viz., $9Q$ and 6.0705σ, are such only when the distribution from which they are computed is strictly normal. The divergence of these values from exact equivalence grows with increasing asymmetry of distribution. Since the validation of test construction requires that the

TABLE I. TYPES OF TRANSVERSIONS FROM ORIGINAL SCORES, STONE NARRATIVE READING SCALE

(Arithmetic mean, 11.15; standard deviation, 3.83)

Score	Deviation	Ratio of Variation	Ratio of Variability	Kental Transversion
20	8.85	0.79	2.31	88.1
19	7.85	0.70	2.04	83.8
18	6.85	0.61	1.78	79.5
17	5.85	0.52	1.52	75.2
16	4.85	0.43	1.26	70.9
15	3.85	0.34	1.00	66.6
14	2.85	0.25	0.74	62.3
13	1.85	0.16	0.48	58.0
12	0.85	0.07	0.22	53.7
11	− .15	−0.01	−0.04	49.4
10	−1.15	−0.10	−0.30	45.1
9	−2.15	−0.19	−0.56	40.8
8	−3.15	−0.28	−0.82	36.5
7	−4.15	−0.37	−1.08	32.2
6	−5.15	−0.46	−1.34	27.8
5	−6.15	−0.55	−1.60	23.5
4	−7.15	−0.64	−1.86	19.2
3	−8.15	−0.73	−2.12	14.9
2	−9.15	−0.82	−2.38	10.6
1	−10.15	−0.91	−2.65	6.3
0	−11.15	−1.00	−2.91	2.0

distribution of measurements from which standardization is effected approach closely to the normal type, the error involved in regarding the $9Q$ and 6.0705σ ranges as exact equivalents is minimized by employing standard measures of dispersion in computing R rather than dispersions computed from data at hand, where skewness of distribution will ordinarily be more marked.

The comparison of the results of measurement in different localities would be greatly facilitated if age and grade norms included both measures of central tendency and measures of dispersion. In my judgment, the arithmetic mean should be the average preferred. It is certainly the most stable of the measures of central tendency. In the use of the formula for the computation of k, if the median is the average which is counted equal to 50, it will be noted that the value of the kental transversion will be contrarily affected by the positive or negative asymmetry of the distribution from which the median is computed. To illustrate, let A be one standardized test and B another. Suppose that the distribution from which the standard median of A is taken is negatively skewed, and that the distribution from which the standard median of B is taken is positively skewed. A moment's reflection will show that the kental values are higher for A than they would have been if the mean instead of the median had been set equal to 50. This is because in a negatively skewed distribution the median is greater than the mean. On the other hand, in the case of B, kental values are less when the median is equated to 50 than is the case when the mean is so equated.

In so far, then, as test makers are interested in the comparability of the results of measurement, they will insist that the arithmetic mean be the average published as norm. As a matter of convenience, however, the median should be published along with the mean. The quartile deviation, because of its greater facility of computation, will doubtless frequently be the measure of dispersion favored. Since the standard deviation is so much employed in experimental work, the publication of both the

standard deviation and the quartile deviation would seem desirable.

The actual transversion of measurements into their kental equivalents would be greatly facilitated by the use of a table constructed like the sample, Table II. In the formula for the computation of k, the variables are the deviations and the

TABLE II. SAMPLE KENTAL TRANSVERSION TABLE

Quartile Deviation	+Deviations					−Deviations		
	0	1	2	3	4	1	2	3
5.0	50	52.2	54.4	56.7	58.9	47.8	45.6	43.3
5.1	50	52.2	54.4	56.5	58.7	47.8	45.6	43.5
5.2	50	52.1	54.3	56.4	58.5	47.9	45.7	43.6
5.3	50	52.1	54.2	56.3	58.4	47.9	45.8	43.7
5.4	50	52.1	54.1	56.2	58.2	47.9	45.9	43.8

measures of dispersions. The table merely relates the values of k to the values of these variables, the quartile deviation being the measure of dispersion employed in the computations. Thus, for a test with a quartile deviation of 5.2, a measurement of deviation +3 would give a kental score of 56.4, and a measurement of deviation —3 would give a score of 43.6 k. Where the standard deviation is the dispersion given for a specified test, it must be multiplied by 0.6745 before the table is used to obtain the transversion. Thus, if the standard deviation of the test is 7.6 and the deviation of a particular score is 2, we multiply 7.6 by 0.6745 and obtain 5.1 as the equivalent quartile deviation. Using the table, we obtain 54.4 as the desired transversion. Inasmuch as the reduction of a standard deviation to its equivalent quartile deviation would need to be done but once for a given distribution of measurements, the operation would add but little to the labor of reading transversions directly from the table.[4]

The formula given for the computation of k may be utilized in transverting teachers' marks into terms which make them directly comparable with standard measurements. This procedure will greatly facilitate measurement in the case of the

[4] The writer is considering the publication of comprehensive tables constructed after the plan of the sample given.

high-school subjects. The steps in the determination of such pedagogical ratings are as follows:

1. Rate all marks according to some determinate standard of attainment. The scale of comparison may be arbitrarily chosen, as 1 to 10, or it may be empirically established by the use of an instrument such as the true-false or completion test.

2. Compute the quartile deviation and median or other measures of dispersion and central tendency.

3. Compute the deviations of scores from the average used and apply the formula for the computation of k.

We may now, I believe, consider that without violating the second of the six demands mentioned at the beginning of this paper—the demand, namely, for simplicity of computation—we have, by the proposed kental method met the first and third demands. In other words, we have provided a unit which will, with considerable exactness, commensurate mental and educational measurements, and we have suggested a method of measurement which will facilitate the comparison of teachers' marks with the results of standard tests. In a subsequent paper we shall see whether or not the kental method is of service in meeting the remaining demands.

THE EFFECT OF ATTENDANCE UPON SCHOOL ACHIEVEMENT

C. W. ODELL
University of Illinois

In connection with a recent experiment[1] the writer had to determine the placement of a large number of pupils. As a basis for so doing, several items of information concerning each pupil were collected. Among these were the scores made upon certain standardized intelligence and achievement tests, the average school marks, and the attendance records. The tests referred to were given near the close of each of the several semesters through which the experiment was continued. In each case the pupil took one or more of the following tests: Pressey Primer Scale, Dearborn Group Tests of Intelligence, National Intelligence Tests, Scale A, Form I, Illinois General Intelligence Scale, Forms I and II, and Myers Mental Measure. At the same time the following achievement tests were used in the appropriate grades: First-Grade Reading Test, Forms A and B (Pressey); Second-Grade Attainment Scale, Forms A and B, (Pressey); Monroe's Standardized Silent Reading Test, Forms I, II, and III; Monroe's General Survey Scale in Arithmetic, Forms I, II, and III.

The point scores made upon these tests were first transmuted into mental or achievement ages and then into intelligence or achievement quotients, as the case might be. The average school mark was given by the classroom teacher and was the average of the marks made in the following seven subjects or in as many of them as the pupil was taking: arithmetic, English, geography, history, reading, spelling, and writing. This mark was for the pupil's work from the beginning of the semester until the time when the tests were given and thus included most of the work of the semester. The attendance record showed the percent of

[1] This experiment is described in Bulletin 12 of the Bureau of Educational Research, University of Illinois.

time that each pupil was present from the first of the semester until the date upon which the tests were given. The attendance of pupils who were not enrolled at the first of the semester was computed on the basis of the time elapsing after they entered.

In determining the placement of pupils their test scores and their average school marks were the weighty factors; their attendance records played only a minor part. In the many conferences with the teachers the author found a rather large number of them who believed that a pupil's attendance record should have a definite influence in determining his placement. Two radically different opinions concerning the extent of this influence were expressed. Some held the belief that pupils who had been absent any considerable portion of a semester should not be promoted at its close. They argued that no matter how bright such pupils might be they should not be promoted until they had covered the work which they had missed. In other words, these teachers believed that all portions of the subject matter included in the course of study should be mastered and that to do so pupils must be present at the time each portion was studied. Furthermore, if the cause of poor attendance was suspected to be a disinclination to attend school these teachers believed that failure was a suitable penalty. They said that if such pupils were promoted the impression would be produced that attendance was not a matter of much importance and thus further truancy would be encouraged.

Other teachers urged the more lenient judgment of pupils with poor attendance records. Their belief was that pupils who had the ability to do satisfactory work should not be held back merely because they were unfortunate enough to be absent due to illness or other causes. This argument, of course, applied only to pupils of average or superior ability. These teachers believed that subject matter which was missed through absence would largely be made up before the pupils left the elementary school. They also believed that failure for the pupil whose absence was due to dislike of school would only result in greater dislike and in more truancy. Because of these two radically divergent view-

points the writer was interested to determine from the available data what effect attendance or non-attendance actually produces upon school work.

A common-sense view of the relation of attendance and achievement is that good attendance will result in better achievement during the period for which both are measured but that attendance during any given period will have comparatively little effect upon the total amount learned up until the end of that time. In order to test the validity of the first of the foregoing statements the increase in achievement age from one testing period to another was found and compared with the attendance record. As the tests were given near the end of each semester this increase in achievement age represented approximately the accretions due to the school work of the semester, in so far as the tests used measured it.

Table I, part A, using grade III-A as an example, is to be interpreted as follows: The coefficient of correlation between the percent of attendance and the increase in achievement age is —0.09, that between attendance and achievement age at the end of the semester is —0.04, that between attendance and achievement quotient at the end of the semester 0.01, and that between attendance and the average school mark 0.14. The first line of part B reads thus: Approximately 75.3 of all pupils were present from 91 to 100 percent of the time. The median increase in achievement age of this group of pupils is 1.4 years; the median achievement age, 11.1 years; the median average school mark, 3.1.

Column 2 of Table I, Part A shows the coefficients of correlation between the percent of attendance and the increase in achievement age by grades, and Column 3 of Part B shows the median increase in achievement age for each of the attendance groups. It will be noted that on the whole there is practically no relation between the percent of attendance and the increase in achievement age. A correlation of 0.01 exists, which really means no correlation at all. It is true that the median increases for the three highest attendance groups seem to show a slight

TABLE I. RELATION OF ATTENDANCE TO ACHIEVEMENT DURING THE SAME SEMESTER

A. COEFFICIENTS OF CORRELATION

Grade	Increase in A. A.	Achievement Age	Achievement Quotient	Average School Mark
I-B[a]				.16
I-A	.02	.08	.06	.09
II-B	−.03	.10	.03	.13
II-A	.02	.09	−.08	.09
III-B	.04	.09	.02	.13
III-A	−.09	−.04	.01	.14
IV-B	.04	.04	.01	.20
IV-A	.03	.08	.07	.22
V-B	−.06	.09	−.02	.10
V-A	.08	.10	−.13	.11
VI-B	.00	.06	−.03	.14
VI-A	.13	.12	.07	.20
VII-B	.03	.10	.00	.18
VII-A	−.01	.03	−.04	.12
VIII-B	−.06	.01	−.10	.12
VIII-A	.04	−.01	−.02	.06
All	.01	.08	.02	.14

B. MEDIANS

Percent of Attendance	Percent of Pupils Included	Increase in A. A.	Achievement Age	Achievement Quotient	Average[b] School Mark
91–100	75.3	1.4	11.1	107.8	3.1
81–90	15.9	1.3	10.6	107.2	2.9
71–80	4.7	1.2	10.4	106.8	2.7
61–70	1.7	1.5	10.4	105.4	2.6
51–60	.8	1.6	10.2	109.2	2.4
41–50	.7	1.6	10.4	113.3	2.5
31–40	.3	.0	9.4	103.3	2.4
21–30	.2	.3	9.2	9.5	2.4
11–20	.2	.5	9.0	103.3	2.5
1–10	.2	1.5	9.9	90.0	2.4
All	100.0	1.4	11.0	107.5	3.0
Total No. of Cases		13,746	23,036	6,874	22,803

[a] No achievement tests were given in grade I-B.

[b] The average school marks are computed on a scale of 5 points, 1 being the poorest and 5 the best. These numerical marks are equivalent to the following expressions: 1, poor; 2, fair; 3, good; 4, excellent; 5, superior.

positive relation, but those for the three attendance groups immediately below are larger than any of the highest three.

The partial correlation between the percent of attendance and the increase in achievement age, with intelligence held constant, was computed. As the correlation between intelligence and increase in achievement age was negative but small, the effect of holding intelligence constant was to increase the apparent correlation of 0.01 to about 0.03—a change too small to signify anything. Thus the partial coefficient of correlation agrees with the other data indicating that there is no connection between attendance and increase in achievement. The evidence, therefore, concerning the relation of attendance to increased achievement does not agree with the common-sense view. According to the data in Columns 3 of Part A and 4 of Part B the achievement ages show a somewhat closer relation. Considering the large number of cases involved, a coefficient of correlation of 0.08 shows a definite, although slight, degree of correspondence. The median achievement ages also tend to show this same relation although they do not increase regularly. The evidence seems to indicate that achievement age, that is to say the total amount of knowledge gained in school, does have a definite, although small, connection with attendance. The achievement quotients show this tendency only slightly, if any, more than the increases in achievement age. The quotients for the four highest attendance groups show small increases from group to group but the others vary considerably. The highest correlation found is in the case of the average school mark.

The medians also show a positive relation between the average school mark and the percent of attendance. In both this case and that of the achievement age the medians seem to indicate a somewhat closer relation to attendance than do the coefficients of correlation. This may probably be accounted for by the fact that, because such a large percent of the pupils falls in the one highest attendance group, the coefficient of correlation does not have the same meaning and validity that it would have were the distribution more nearly normal. The median

percents of attendance for all pupils of each given achievement age, quotient, increase, and school mark were also computed but are not presented in this article. These medians show the same tendencies that have been mentioned, namely, that groups of pupils who have the higher achievement ages and school marks tend to be those who also have the higher percents of attendance. There is practically no such tendency in the case of the achievement quotients and increases in achievement age. The evidence presented in Table I seems to show that there is some relation between the attendance and the achievement during the period for which the attendance is computed.

It was also thought worth while to attempt to ascertain the relation of attendance during one semester to quality and quantity of school work during the next. Table II presents data similar to those in certain of the columns of Table I except that they show the relation just referred to. These data do not show any tendencies strikingly different from those previously presented. Whereas, for all grades, the coefficient of correlation between increase in achievement age during the same semester and attendance is 0.01, it is 0.00 for the succeeding semester. Both of these figures show that there is no correlation. The correlation of the average school mark with attendance, which is 0.14 for the same semester, is 0.08 for the next, which also shows a slight correspondence. The median increases in achievement age do not show even the slight tendency toward a positive relation which is shown by those presented in Table I. Medians for the average school marks, however, show that in the five groups having the highest attendance there is some relation. They do not, however, show quite as pronounced a relation as those for the same semester.

Probably the most noteworthy fact shown by Table II is that the median increases in achievement age for the three lowest attendance groups are very high. This may be due to the fact that these groups contain such a small number of pupils that the medians are not reliable. If these medians are reliable it may be that the pupils in these groups made such small increases

in achievement age during the semester for which the attendance was taken that they were under special pressure during the succeeding semester. This pressure may have been on the part of the teacher or due to hard effort and diligent application on their own part or both.

Summing up the data, the following are the chief inferences to be drawn: (1) The percent of time which pupils attend school

TABLE II. RELATION OF ATTENDANCE TO ACHIEVEMENT DURING THE FOLLOWING SEMESTER

	A. Coefficients of Correlation		B. Medians			
Grade	Increase in A. A.	Average School Mark	Percent of Attendance	Percent of Pupils Included	Increase in A. A.	Average School Mark
I–B.	...	.10	91—100...	75.3	1.3	3.1
I–A.	−.13	.12	81— 90 ..	15.9	1.5	3.0
II–B.	.06	.02	71— 80...	4.7	1.0	2.9
II–A.	−.07	−.01	61— 70...	1.7	1.1	2.9
III–B.	−.03	.08	51— 60...	.8	.8	2.6
III–A.	.04	.11	41— 50...	.7	.8	2.9
IV–B.	.07	.08	31— 40...	.3	1.9	2.7
IV–A.	.04	.08	21— 30...	.2	3.0	3.0
V–B.	−.03	.11	11— 20...	.2	2.5	2.9
V–A.	.02	.08	1— 10...	.2	3.3	3.1
VI–B.	.03	.04				
VI–A.	−.05	.06				
VII–B.	.00	.11				
VII–A.	.02	.06				
VIII–B.	.07	.17				
VIII–A.	(−.55)*	.11				
All....	.00	.08	All....	100.0	1.3	3.1
			Total No. of Cases..	...	7,074	12,773

* The figures in the parentheses in this and the succeeding tables are so placed because they are based upon such a small number of pupils as to be very unreliable.

has a small but definite positive relation to their achievement age at the end of the period for which the attendance is computed. (2) It has practically no effect upon their actual increases in achievement age except in the case of those pupils whose attendance records are very poor. In such cases the increase is small for the same semester and large for the succeeding one. (3) Attendance is most closely related to the pupils' average school marks. This is especially true when the marks for the

same semester are considered and somewhat less so when those for the following semester are taken. The percent of time which pupils attend school appears to condition their achievement but not to the large extent that many have supposed.

There are two or three other questions of some interest connected with the attendance and non-attendance of pupils at school. One of these is whether or not there is any relation between the intelligence of pupils and their attendance. To secure some data upon this question a brief study was made of the records of the pupils who were in school during the four semesters for which data were obtained. The average intelligence quotient resulting from four mental tests[2] was correlated with the average attendance for the four semesters. Table III presents the coefficients of correlation and medians obtained.

TABLE III. RELATION OF ATTENDANCE TO INTELLIGENCE QUOTIENT

Grades	Coefficients of Correlation	Percent of Attendance	Median Intelligence Quotient
I-B	−.87	1- 10	....
I-A	−.11	11- 20	....
II-B	.05	21- 30	....
II-A	.06	31- 40	(95.0)*
III-B	.14	41- 50	(80.0)
III-A	−.08	51- 60	(100.0)
IV-B	.09	61- 70	100.0
IV-A	.00	71- 80	99.1
V-B	.11	81- 90	99.9
V-A	.13	91-100	103.2
VI-B	.02		
VI-A	.13		
VII-B	.16		
VII-A	.05		
VIII-B	.04		
VIII-A	.00		
All	.06	All	102.5
		Total number of cases	3,824

* See footnote *a* Table II.

[2] Although each pupil was tested four times he was not tested with four different mental tests. In some cases pupils were tested with two different mental tests, one of which (Illinois or Pressey) was used three times. In other cases three different tests were used, one of them (Illinois or Pressey) being used twice.

The correlation for all pupils was 0.06, which shows a very slight positive relation. The median I.Q. for the highest attendance group is distinctly above that for any of the others but with this exception the medians show practically no tendency toward a positive relation. There is also a very slight tendency for the median percents of attendance for the groups of pupils having different intelligence quotients to increase along with the latter. On the whole, the evidence justifies the conclusion that the relation between intelligence and school attendance is very slight. It is even possible that this slight relation may be explained as follows: It is well known that our intelligence tests do not test native intelligence alone but that the scores made thereon are influenced by training both in school and elsewhere. It is, therefore, possible that the higher median intelligence quotient of the group having from 91 to 100 percent of attendance is due to the fact that because of better school attendance this group is able to make higher scores upon the intelligence tests.

Another matter of interest is whether or not some pupils have practically perfect attendance records semester after semester and others poor records due either to confirmed truancy or continued poor health. In other words, are most of the pupils who attend rather regularly any one semester likely to do so the succeeding semester? In order to secure data upon this point the attendance records during the first and second semesters of the same school year were correlated. Table IV presents the coefficients and medians obtained. With the exception of grade VIII-A, in which too few pupils are concerned to make the figures reliable, it will be noted that the correlation is not above 0.32 in any grade, while that for all the grades is only 0.19. This indicates that there is a small but definite amount of correlation between attendance during successive semesters. The medians tend to show the same thing. For the six groups having the best attendance records during the first semester there is a corresponding increase in median attendance for the second. A similar condition is found when the comparison is reversed and the attendance groups for the second semester are compared with the median

TABLE IV. RELATION OF ATTENDANCE DURING TWO SUCCESSIVE SEMESTERS

Grade	Coefficient of Correlation	First Semester Percent of Attendance	Percent of Attendance Second Semester (Medians)
I–B	.15	1– 10	91.6
I–A	.12	11– 20	91.0
II–B	.07	21– 30	91.8
II–A	.27	31– 40	91.6
III–B	.28	41– 50	87.1
III–A	.21	51– 60	89.1
IV–B	.19	61– 70	91.0
IV–A	.05	71– 80	92.0
V–B	.08	81– 90	92.6
V–A	.08	91–100	94.6
VI–B	.24		
VI–A	.18		
VII–B	.28		
VII–A	.22		
VIII–B	.32		
VIII–A	(.70)[a]		
All	.19	All	94.2
		Total number of cases	5,637

[a] See footnote a in Table II.

attendance of the same pupils for the first. The slight degree of positive relationship between attendance during a semester and the one succeeding it shown by these data is much less than many school officials and teachers have believed to exist. It has often been assumed that poor attendance is largely due to persistent truancy semester after semester. The figures given above do not support this view. It should be noted, however, that these data were secured from certain schools in Chicago and that in the Chicago public school system the compulsory attendance department is distinctly effective. Pupils, therefore, who if left to themselves and their parents would be absent rather consistently are kept in school.

It is probable that some readers of this article will wish to compare the attendance records of their pupils with those of the pupils concerned in this experiment. One way of doing so is by means of the figures given in the second columns of Tables I and II, parts B. These columns show that almost exactly

three-fourths of the pupils attended from 91 to 100 percent of the time, less than one-sixth from 81 to 90 percent, and so on down to one in 500 attending 10 percent or less. The median attendance for all pupils concerned was 94.4 percent and the mean or average attendance was 90.6 percent. In other words, the median pupil was absent about one week per semester and the average pupil almost two weeks. It was also found that although considering the semesters separately about 2.2 percent of the pupils attended less than 61 percent of the time, when the average attendance of the pupils in school during the whole four semesters was computed only about 0.2 percent had attended less than this amount of time. This agrees with what might have been expected from the fact that the correlation of attendance in successive semesters was found to be low.

A GRAPHIC RATING SCALE FOR TEACHERS

MAX FREYD
University of Pennsylvania

Recently the writer[1] listed the outstanding methods of rating and the various lines of approach which had been followed in attempting to compare and evaluate these methods. It is impossible to repeat here the substance of that article, but the broad conclusions to which the writer has been led are the following:

1. That it is next to impossible to make statistical comparisons between kinds of rating scales. It seems likely that ratings could be evaluated by correlating them with some variable, such as an objective measure of the particular trait upon which ratings are made. But ratings are ultimates in psychology, and cannot be evaluated with reference to a known criterion. If such a criterion were available there would be no need to make ratings. Ratings themselves have often been used as criteria for finding out just what trait is measured by a psychological test. The remaining statistical methods are all subject to errors of interpretation, and as yet no figures indicate conclusively that one type of rating scale is superior to another. The comparison of rating scales on non-statistical bases, such as ease of administration and scoring and simplicity of the rating task, is probably as far as we can now go with profit.

2. That at present a graphic rating scale promises to be popular, due to its concreteness, simplicity, universality, and ease of administering and scoring.

The present article will be devoted to a practical application of the graphic rating method to the rating of teachers. The number of rating scales for teachers is constantly increasing, and an addition must carry with it some justification. Before proceeding to this justification, however, it will be necessary to present the rating plan.

[1] Freyd, Max. "The Graphic Rating Scale." *The Journal of Educational Psychology*. 14:83-102. February, 1923.

This plan covers seventeen traits, for each of which a scale is provided. The traits are as follows:

1. Physique and bearing.
2. Neatness in dress.
3. Self-possession.
4. Sociability.
5. Physical energy.
6. Alertness.
7. Sense of humor.
8. Self-assertion.
9. Tact.
10. Popularity.
11. Impartiality.
12. Patience.
13. Acceptance of criticism.
14. Ability to think on feet.
15. Fluency of speech.
16. Distinctness of speech.
17. Interest in teaching.

Notice that the scale does not concern itself with efficiency in teaching. It furnishes a means for measuring the psychological characteristics which are assumed to be fundamental to the acquisition of a successful teaching technique. The traits which it covers are for the most part components of that complex known as personality, which at present cannot be measured objectively. It is not to be implied that these traits are totally innate or insusceptible of improvement. Individuals will no doubt disagree with this analysis and find grave omissions or errors of inclusion, but no one can dispute the fact that in the main it covers the abilities essential to successful teaching which can be or should be rated. Some items are omitted, such as health, freedom from family disturbances, etc., which may be fully as essential as any in the list, but they must be approached in other ways. The analysis upon which the scale is based was the result of almost a semester's work by a class of students of which the writer had charge.

The directions for making the ratings and the scales for the first five traits are given below.

DIRECTIONS FOR USING THE RATING SCALE

1. Let these ratings represent your own judgments. Please do not consult any one in making them.

2. In rating this person on a particular trait, disregard every other trait but that one. Many ratings are rendered valueless because the rater allows himself to be influenced by a generally favorable or unfavorable impression which he has formed of the person rated.

3. When you have satisfied yourself on the standing of this person in the trait on which you are rating him, place a cross (X) at the appropriate point on the horizontal line. You do not have to place your cross directly above a descriptive phrase. These phrases merely represent steps on a *continuous* scale. *You may place your cross at any point on the line.*

4. Do not write on the dotted lines. If you have any comments, write them at the margins or at the end of this blank.

5. It will be found advisable to read the whole blank before commencing to rate, in order to realize the scope of the blank and to differentiate the traits.

1. How does he impress people by his physique and bearing?

Excites admiration. Very impressive.	Noticeable for good physique and bearing.	Makes satisfactory impression.	Unimpressive physique and bearing.	Arouses repulsion. Looked down on.

2. Does he appear slovenly or neat in his dress?

Slovenly and unkempt.	Somewhat indifferent and careless in dress.	Unnoticeable in dress.	Noticeably neat and clean.	Fastidious. Almost a dude.

3. Is he self-conscious or self-possessed?

Painfully self-conscious and ill at ease.	Frequently embarrassed or flustered.	Self-conscious at times.	Usually unmoved by actions or remarks with reference to himself.	Always at ease. Self-possessed.

4. What is his social attitude?

Extremely breezy and informal.	Cordial and congenial.	Meets one half way.	Slightly reserved.	Constrained and formal.

5. How much physical energy does he put forth?

Listless and sluggish.	Slack and half-hearted.	Gets things done.	Active.	Full of pep and energy.

Rating is done very simply by making a check along each line. It thus provides the answer to the question which introduces each scale. The phrases beneath the rating line make the various degrees of the trait more concrete. Rating is thus equivalent to describing a person in short phrases (with some interpolation), but at the same time the ratings may be treated numerically, as will be shown later on.

The scales recognize the fact that both extremes of a trait *may* be undesirable. Ordinarily we rate people on "neatness"

or "sociability," assuming that a high standing represents a desirable condition. But neatness may be carried to a point where it is no longer a virtue, and there are individuals in whom social informality is a vice. The second and fourth scales recognize these possibilities. In most of the scales the central phrase is a neutral one, and this also may often represent an undesirable condition as regards the trait. The standing of an individual in any trait must be interpreted together with his standing in the other traits, in order to observe what compensation of abilities is present.

Another noteworthy point is that the scales are so arranged that the more desirable extremes of the traits are alternated. In the first scale, for example, it will be conceded that the left-hand extreme is the more desirable, in the second we are in doubt, in the third the right-hand extreme is the more desirable, and so on. This arrangement is made to reduce the tendency for a rater to check at one side of the page, an error which is based upon the formation of a favorable or unfavorable attitude toward the person who is rated. If the favorable ends of all scales were to the right, the rater would be more likely to allow himself to be influenced in his ratings on each trait by a general attitude toward the person rated, especially so since it is easy to fall into the habit of checking at one side of the page.

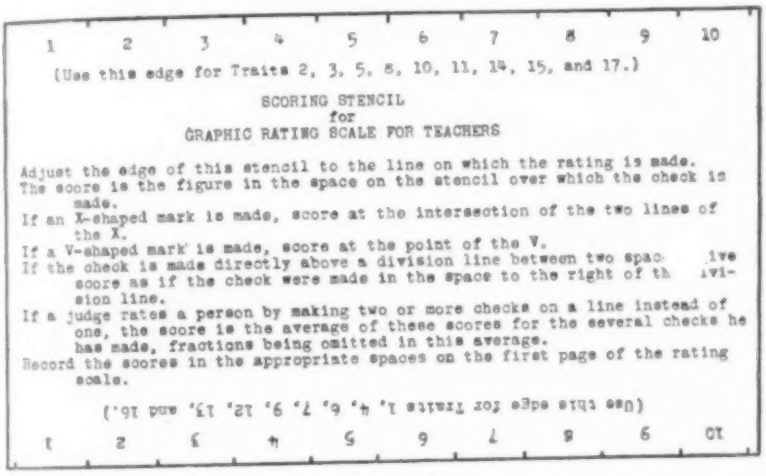
A complete rating on any one person can be made in a very brief period, five or ten minutes usually being sufficient.

If the rating is merely for the purpose of forming a general qualitative impression of the individual, no steps are required beyond simply observing where the checks are made. But if one person rated is to be compared with another, if one rater is to be compared with another, or if ratings are to be made on the same person over a period of time to note what improvement takes place, then some numerical expression of the rating is necessary. This can be obtained with the use of the scoring stencil which is reproduced in Figure 1:

The two marked-off spaces at the edges of the stencil correspond in length with the lines on which the ratings are made.

These portions of the edge of the stencil are placed directly beneath each line on which a rating has been made, and the score is read off as indicated on the stencil. These scores are then recorded in spaces provided for that purpose on the first page of the rating scale.

Before the stencil was constructed, one end of each scale was arbitrarily judged to be the more desirable extreme of the trait. Since it would simplify matters to have a high score always indicate a high or desirable amount of the trait, two scoring edges of the stencil were found necessary, one with scores reading from right to left and one with scores reading from left to


1 2 3 4 5 6 7 8 9 10

(Use this edge for Traits 2, 3, 5, 8, 10, 11, 14, 15, and 17.)

SCORING STENCIL
for
GRAPHIC RATING SCALE FOR TEACHERS

Adjust the edge of this stencil to the line on which the rating is made.
The score is the figure in the space on the stencil over which the check is made.
If an X-shaped mark is made, score at the intersection of the two lines of the X.
If a V-shaped mark is made, score at the point of the V.
If the check is made directly above a division line between two spaces, give the score as if the check were made in the space to the right of the division line.
If a judge rates a person by making two or more checks on a line instead of one, the score is the average of these scores for the several checks he has made, fractions being omitted in this average.
Record the scores in the appropriate spaces on the first page of the rating scale.

(Use this edge for Traits 1, 4, 6, 7, 9, 12, 13, and 16.)

10 9 8 7 6 5 4 3 2 1

FIGURE 1

right. This has been taken care of with the present scoring stencil, and the scales for which each edge is to be used are indicated on the stencil near that edge. High scores in general indicate desirable amounts of the traits.

The scales could be scored from 1 to 5, or from 1 to 20, but the calibration from 1 to 10 has been adopted as the most satisfactory.

Now, it may be maintained with justice that a high score does not necessarily indicate that the person will be a successful

teacher. It may be found that in a certain scale a score of 4 or 5 is much more indicative of success than a score of 10. The most favorable score will vary with the locality and type of teaching that is demanded, and no norms may be set which have general validity. Each institution using the scale will have to decide for itself experimentally which is the most desirable amount of each trait, by rating its teachers and comparing the scores made by the best and the poorest teachers. It may be possible to use the stencil as it is and then to interpret the scores by placing in a column adjacent to the list of raw scores in the various traits, a list of equivalent scores in terms of the standards which have been set for the given situation. Or it may be necessary to construct new scoring stencils in which for some traits the zero point is not at one end of the line but at the center, and from which scores may then be read in both directions. It must always be remembered that each trait is to be interpreted in relation to the other traits. This interpretation is most easily understood when placed in the form of a total score for all traits as obtained by the scoring stencil here described.

It may also be maintained that some of the traits are more important as regards predicting success than other traits. For example, much greater importance may attach to interest in teaching than to neatness in dress. Even if true this is no shortcoming in the scale. When the relative importance of the various traits has been determined statistically, the raw scores may be weighted in proportion to the importance of the trait.

Two problems thus follow the recording of scores on the front page of the blank: The transmutation of scores so that a high score will always be predictive of success in teaching, and the assignment of weights to the various traits. Research must solve these problems.

We may now point out the salient features of this type of rating scale. Whether or not they constitute advantages over other types of rating scales must be left to experimentation.

1. The scale does not cover the technique of teaching but rather the psychological traits which probably underlie success in teaching.

2. The scale is interest compelling. It does not fail to draw the interest of the rater. This is a very important point if ratings are desired from individuals at a distance who lack motivation and who ordinarily would neglect to return the scale completely filled out.

3. By occasionally reversing the scales so that the most desirable extreme does not always lie at one edge of the sheet, and by recognizing that the highest score does not *necessarily* indicate the most desirable amount of the trait, the scale has a tendency to reduce the amount of "halo," or tendency to judge an individual high in everything or low in everything.

4. The scale is concrete; the rater has a cue to go by. He does not have to rate a person 1, 2, 3, 4, or 5 in "neatness," but rates the person in terms of phrases which he might ordinarily apply to him in conversation.

5. The rater does not have to make a quantitative judgment. This quantitative interpretation is made later by the person who scores the blank.

6. The scale may be filled out very quickly.

7. The resulting scores may be as fine as desired.

8. The scale is universal. No master scale is required for each new group that is rated.

9. The scale permits comparable ratings without requiring the rater to know all the members of a group. In many other types of rating scales, statistical equalization of scores must be made if some of the raters do not know all the members of the group that is being rated.

This type of scale should find use not only in rating for selection and placement, but also for self-rating and self-improvement; or more broadly stated, on any occasion requiring the systematic recording of personal judgments.

THE OTIS CORRELATION CHART

ARTHUR S. OTIS
Director, Test Service, World Book Company, Yonkers, New York

Since the Otis Correlation Chart, recently published,[1] may appear to some rather complicated, it seems desirable to make a few brief explanations as to the purpose of the various provisions that have been incorporated in it.

It is perhaps a mistake to claim simplicity for this chart. It might be more correct to say that the aim has been to provide a chart which would make for greater ease and accuracy of computation. I shall attempt to explain in this article in what sense it is felt that by the use of this chart a coefficient may be calculated with greater ease and accuracy. I shall say here merely that "ease" is used in reference to the amount of mental effort required; and that in this connection provision is made so that every step is performed in the right order and none omitted.

The feature of the Otis Correlation Chart that makes for this ease is the complete job analysis whereby every step necessary in arriving at the coefficient of correlation is laid out successively. One merely has to take one step at a time, following minute directions without any forethought or hindsight. Indeed, it has been found that an intelligent person can actually work a correlation on this chart without ever having seen a correlation formula before.

I shall mention first some general considerations regarding the calculation of a coefficient of correlation, and then take up seriatim the various provisions that have been introduced in this chart, explaining the reasons for each.

One of the very earliest methods of finding a correlation consisted of the following steps: (1) expressing the values of the X-variable in terms of deviations from the mean of the X-variable (calling these deviations x's), and similarly expressing the values of the Y-variables in terms of deviations from the

[1] Published with complete directions for its use by World Book Company, Yonkers-on-Hudson, N. Y., 1922.

mean of the Y-variables (calling these deviations y's); (2) finding the product, xy, of each x and its corresponding y and finding the algebraic sum (called Σxy) of these products; (3) finding the sum of the squares of the x's (called Σx^2), and similarly, finding the sum of the squares of the y's (called Σy^2); (4) solving the formula (r is the correlation coefficient),

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \, \Sigma y^2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

There are various objections to the use of this method. First, it is impossible to tell from an inspection of two columns of variables whether the relationship between them is rectilinear or curvilinear. Moreover, it is not easy to detect the presence of extreme cases lying outside the general scope of the scatter diagram and which require reference to the original test papers in order to correct any possible errors in scoring. These are two reasons why it is necessary to plot a series of paired variables in a scatter diagram such as shown in the plot in Figure 1 in order to inspect carefully the scatter of the plotted points and to correct any errors before calculating the coefficient. The present correlation chart therefore rests first of all upon the assumption that a scatter diagram should be made in every case before calculating a coefficient of correlation.

A second objection to the use of Formula (1) is the rather tedious process of subtracting the mean of each variable from every value of that variable and squaring these differences. In the first place, the difference will, in practically every case, involve a fraction. If this fraction is disregarded an error is introduced, and if it is not disregarded, the computation becomes extremely tedious. The present correlation chart obviates this difficulty by an enlargement of the correlation formula in such a way that by means of a small amount of additional calculation requiring but a very few minutes, there is applied in the calculation of the coefficient a correction that obviates the necessity of dealing in any way with fractions and that yields the same accuracy as though fractions were taken into account. This will be explained more fully further on.

Correlation between Higher Exam. Form Q. (x) and Higher Exam. Form B. (y) r_{xy} = .917 ± .009

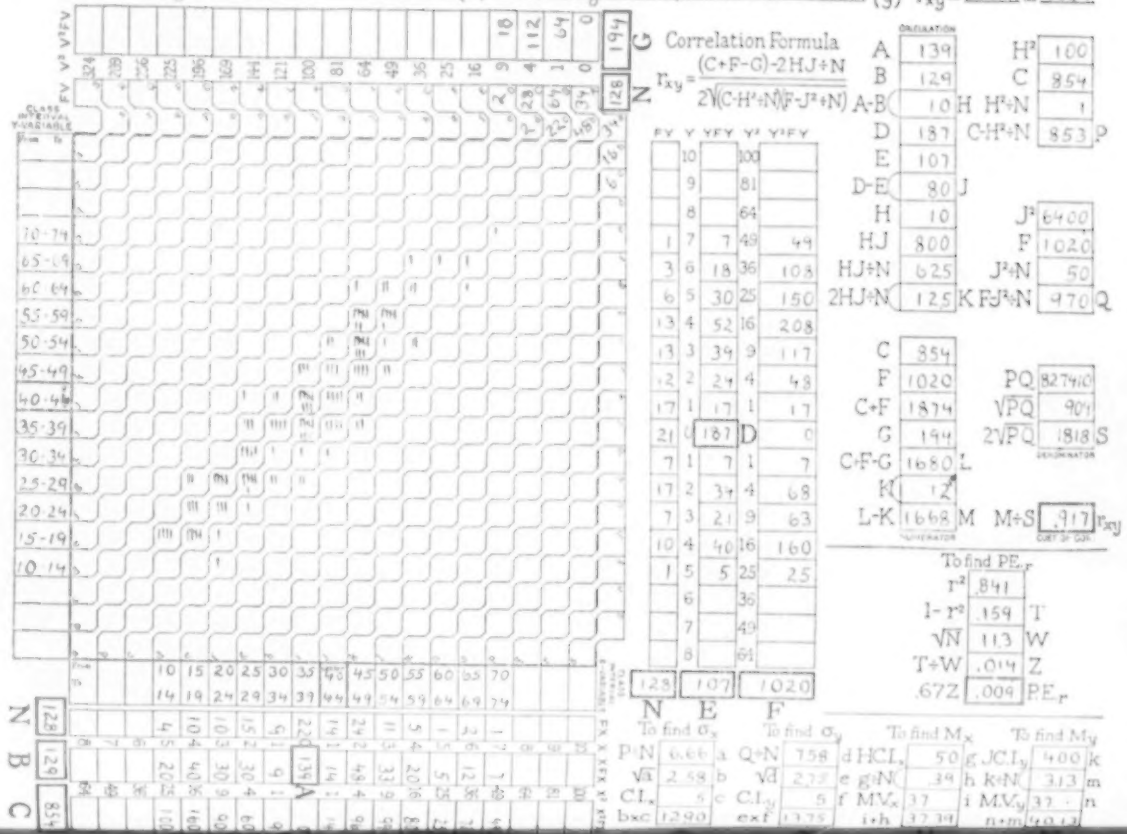


FIGURE 1. THE OTIS CORRELATION CHART

A third drawback to the use of Formula (1) is that the numbers that must be dealt with are unnecessarily large. For example if x equals 127 and y equals 159, it is necessary to multiply these two rather large numbers together and to add the product to other products of similar magnitude. This difficulty can be overcome to a considerable extent by assigning arbitrary scales in place of the actual scales so that, for example, values of x from minus 2 to plus 2 are called 0, those from 3 to 7 are called 1, those from 8 to 12 are called 2, etc., thus condensing five units into one. This procedure is facilitated by the use of a scatter diagram—a fact which constitutes an additional reason for drawing such a diagram.

A fourth difficulty in using Formula (1) is the necessity for calculating each product of the paired variables separately. This may be overcome to some extent if the variables are plotted in a scatter diagram as shown in Figure 1 and considered in groups. Thus, if there are seven cases in which y equals plus 4 and x equals plus 2, the sum of the products of x and y in these seven cases is simply 7 times the product of 2 and 4. That is, the product is found for one pair of variables and this product is multiplied by the number of variables whose points fall together in the same compartment of the scatter diagram. Even this procedure, however, does not greatly simplify matters.

In the present correlation chart all finding of products of paired variables is eliminated by the use of a formula that substitutes for this procedure simply the counting of the points falling in each diagonal row of compartments and multiplying this sum by the square of the difference between X and Y represented by that diagonal row.

For example, in the longest diagonal row extending from the lower left-hand corner to the upper right-hand corner of the square plot, each value of X is equal to the corresponding value of Y. (The values of X and Y referred to here are the arbitrary scales appearing under the letters X and Y at the bottom and right side of the square plot.) For each compartment on this diagonal the difference between X and Y is therefore zero.

Accordingly, the number of points along this diagonal row of compartments is multiplied by zero.

Each point in the diagonal rows immediately above and immediately below this one (those labeled b at the lower end) is such that the difference between X and Y is either plus 1 or minus 1, and in either case the square of this difference is plus 1. Therefore the number of points in these two rows is simply multiplied by 1. The number of points in the diagonal rows next above and next below these diagonals—i.e., the number of points in the diagonals marked c—is multiplied by 4, etc., and the sum of these products is found in the box lettered "G" on the correlation chart.

This procedure has been found to require much less time and much less mental effort than the rather tedious calculations of the products of x and y for each compartment, multiplying each by the number of cases in that compartment and finding the sum of these products.

Whenever the points in a scatter diagram in the Otis Correlation Chart are so placed that the mean of the X values corresponds exactly with the middle value of the interval opposite letter A, and are so placed that the mean of the values of the Y variable corresponds exactly with the middle value of the interval opposite the letter D, then the coefficient of correlation might be found by means of the formula

$$r = \frac{C + F - G}{2\sqrt{C\,F}} \qquad (2)$$

in which C equals Σx^2, F equals Σy^2, and G equals $\Sigma(y-x)^2$, and in which the values of C, F and G are found as explained in the directions accompanying the chart.

Having once found the values of C, F, and G, the coefficient of correlation could be calculated with five additional operations.

If we let $\Sigma xy = G'$, $\Sigma x^2 = C$, and $\Sigma y^2 = F$, then Formula (1) can be expressed thus:

$$r = \frac{G'}{\sqrt{C\,F}} \qquad (3)$$

Now considering that C and F and the square root of their product have to be found both in Formula (2) and in Formula (3); that when the values of C and F are found for the denominator, they can be inserted in the numerator of Formula (2) by simply copying them; and that the value of G can be found in less than half the time it takes to find the value of G′, we believe that the basic formula used in the chart is really not appreciably more complicated than the traditional formula.

We have now to consider the corrections to be applied to Formula (2) when the means of the distributions do not fall in the middle of the chart—the corrections in which the value $\frac{2HJ}{N}$ is subtracted in the numerator, and in which $\frac{H^2}{N}$ is subtracted from C and $\frac{J^2}{N}$ from F in the denominator. These corrections are identical with the corrections necessary to apply to the numerator and denominator of the traditional Formula (3) in order to correct the coefficient for the differences between the assumed means and the true means of the two variables.

If we apply these same corrections to Formula (3), the traditional formula will appear as follows:

$$r = \frac{G' - HJ \div N}{\sqrt{(C - H^2 \div N)(F - J^2 \div N)}} \quad \ldots\ldots\ldots (4)$$

This is about as complicated as the formula used in the chart. Now, if it is desired to calculate the exact coefficient—with corrections made for assumed means—it is necessary to go to the labor of solving one of these complicated formulas. The advantage of the present chart is that each step of this process is laid out in such a way that the complicated formula can be solved even without referring to it. Indeed the advantage of the chart, as has been said, is that it is not even necessary to know how to solve such a formula in order to obtain the coefficient.

The correction introduced in Formula (4) and in the formula used in the chart is now generally accepted by statisticians as being necessary for ordinary accuracy in calculating a coefficient,

and as being a decided advantage over the cumbersome method of dealing with fractional deviations.

In the present correlation chart the zero points of the arbitrary X and Y scales at the bottom and right side of the square plot are so placed as to be approximately opposite the center of the plot. There is a distinct advantage in this arrangement, since values of X and Y are kept small. For example, it is impossible for the value of X or Y to exceed 10 on this chart, whereas if the zero points of these scales were placed at the bottom and left side of the square plot, the values of X and Y could run as high as 18, with consequently very large products in the Y^2FY and X^2FX columns. In the present correlation chart the numbers are so small that these products can be calculated mentally.

Negative values of the correction are avoided in most cases, as explained in the directions accompanying the chart, by placing the zero points of the arbitrary scales a little off center.

To the layman reading the directions for the first time, the handling of negative numbers, when these occur, seems rather confusing; but it is not difficult to learn exactly what to do in this case, and the advantage of the small numbers in correlation far outweighs the slight difficulty in attending to negative numbers now and then.

The complicated appearance of the chart is increased by the provisions for finding the probable error of a coefficient and for finding the standard deviations and means of the x and y distributions. But it is not necessary to calculate any of these auxiliary measures. If they are needed, the provision made for their calculation on the chart greatly assists in finding them, since the method of their calculation is analyzed and laid out in detail, step by step; and they can be calculated without cudgeling the memory for the formula and without fear lest some step be forgotten.

The use of letters in place of the more complicated and technical mathematical symbols assists the layman in following the steps in the calculations, and the use of small letters[2] around

[2] I am indebted to Dr. Herbert A. Toops for this idea.

the edges of the correlation plot is a distinct aid in keeping track of the diagonal columns and in adding points.

The use of the little grooves joining the diagonal rows of compartments (a device originated by the author) has been found a distinct aid in following the diagonal rows of compartments.

There is furnished on the back of each chart a table of products[3] in which the numbers from 1 to 40 are multiplied by 2 and 4, 3 and 9, etc., up to 18 and 324. A special feature of this table is its arrangement so that it may be applied, for example, to the distribution of frequencies of Y values in such a way that the YFY and Y^2FY products may be read directly from the table opposite these frequencies. (The same is true for the X and V distributions.) This enables the user to find the values of A, B, C, D, E, F, and G without doing any multiplying whatever.[4] This table has two distinct advantages. It relieves the mind of the effort of making these multiplications, since the products can be read directly from the table, and it increases accuracy.

Summary

The advantages in the calculation of a coefficient of correlation which are had by virtue of the various provisions of the Otis Correlation Chart are as follows:

1. Observation of the character and the relationship between variables seen from a graphic representation of the correlation.

2. The possibility of checking up errors in scoring by the extreme position of points in the correlation plot.

3. The use of small numbers in calculating; first by grouping the values of the variables in intervals assigned consecutive arbitrary values of X and Y, and second because of the assigning of the zero points in these arbitrary scales near the middle of the plot.

[3] For this idea I am also indebted to Dr. Toops.

[4] Although the numbers in the table of products representing the frequencies do not run above 40, it is possible to use the table for frequencies of any size. Thus, suppose it is desired to find 67 times 49. This may be done by finding 40 times 49 and 27 times 49 and adding these products.

4. The substitution of the simple process of counting points in a diagonal row of compartments for the more tedious calculation of products.

5. The use of single letters in place of more complicated mathematical expressions.

6. The complete job analysis of the calculation of a coefficient of correlation to relieve the mind of the burden of interpreting a correlation formula and to avoid the possibility of error through neglecting some step.

7. The use of grooved diagonals to aid the eye in following a diagonal line of compartments.

8. The use of tables to obviate a great deal of mental effort in calculating products.

It is no easy matter to calculate a coefficient of correlation under any circumstances. Neither can a coefficient be calculated in a very short time if the number of cases is very large and all necessary precautions for accuracy are taken. The most important considerations in the construction of a chart are:

That accuracy be insured in overcoming the possibility of the omission of any steps, by means of an orderly arrangement of these steps with operations indicated.

That the formula used as the basis of the chart be a complete and accurate formula and not a mere approximation.

That the chart be so constructed and labeled as to help the eye in following columns and making additions.

That the directions accompanying the chart be entirely unambiguous and such that with careful study and attention to detail an intelligent person can proceed to carry them out and succeed in finding the coefficient without reference to a textbook.

It is with these considerations in mind that the Otis Correlation Chart has been constructed.

Editorials

WHAT IS THAT IN THINE HAND?

A few years ago Superintendent Joseph S. Taylor wrote an article entitled, "The School's Survey from the Superintendent's Point of View." He made effective use of a portion of the biblical account of God's choosing of Moses to lead the children of Israel out of Egypt. In particular he referred to these words:

> And the Lord said unto him, What is that in thine hand? And he said, A rod.
>
> And He said, Cast it on the ground. And he cast it on the ground, and it became a serpent. (Exodus 4:2, 3.)

As we recall Superintendent Taylor's treatment of this subject, he regarded the survey as essentially an answer to, "What is that in thine hand?" As such, the survey is not an end but a means. It helps to define problems and to obtain the facts about problems. It does not, except incidentally, solve problems.

In the ancient narrative the rod of Moses represented the ordinary ability of the plain man—the plain man which Moses thought himself to be. Touched by divine providence the rod became a mighty weapon. It appears and reappears in the old Hebraic account. It not only became a serpent when cast upon the ground, thus symbolizing its miraculous power, but, smote upon the rock, it caused the waters to rush forth; stretched over the land, it caused the pestilence of frogs to appear; extended toward Heaven, it caused fire and thunder and hail to be precipitated. The rod of Moses became a symbol of the great lawgiver's divine power and of his divine election.

To a teacher the answer to the question, "What is that in thine hand?" is a true statement of the facts of the teaching processes and of the school situation as it enters into his daily

experience. It is not opinion, nor guess work. It is the material upon which vital decisions may be based. Are girls better readers than boys? Are two groups in arithmetic better than one? Can a class of forty be taught as effectively as a class of twenty? Is this textbook in geography better than that? Does the socialized recitation get one anywhere? Such questions are not to be answered by mere expressions of opinion. The evidence is in the hand of the teacher if he will use it. It may not be complete but it is profoundly suggestive—a veritable rod of Moses capable of unsuspected uses.

The survey idea does not need to be embodied in a pretentious report. It can be present in the classroom; and indeed, if it is not present there, its existence elsewhere will be of little consequence.

Most teachers do not realize what they hold in their hands. Rightly conceived they are engaging from day to day in the most fascinating of adventures. There are no contacts in all the world so interesting as contacts with human personalities. The teacher has not only contact but continuous communion with large numbers of personalities, each exhibiting peculiar attributes, each throwing new light upon old notions. The teacher who appreciates these things will record his experience. He will not trust to the judgment of the moment, for to do so will be to invite error. We have all realized, for example, how strongly our judgments are effected by the recency of our experience. The fact that we have lately observed a certain instance colors our thinking far more than the occurrence of an opposite instance some time ago. Yet in their bearing on the question which may be at issue these instances may be equally important. There is no better antidote to the very human tendency to be more greatly impressed by our most recent experiences than that impersonal record-making which places all observed facts on an equal footing in this respect. It is when facts are carefully recorded and competently analyzed that real knowledge, real expertness, and real progress in education become possible.

Let us not, therefore, belittle or misprize the work we are doing. Let us record the facts connected with it in order that we may have a better understanding of it. In every classroom and in every office there is material for a local survey. We can, if we will, know the truth and use it to accomplish better work than we have ever accomplished. It is the answer to the question, "What is that in thine hand?"

B. R. B.

PROFESSIONAL EDUCATION IN COLLEGE AND UNIVERSITY

There seems to be a dawning suspicion among college teachers that professional training, apart from training in subject matter, may be important. There are intimations, obscure and infrequent, but nevertheless discernible, that the educational man in coming into his own among his colleagues in higher education. Typically the professor outside of the department or college of education regards the educational faculty as a mistake, and the money spent on the department of education as a diversion of funds from more worthy uses. He thinks teacher training is a farce and denies that education affords a field either of teaching or research. But the typical college professor is not having everything his own way. There is defection in the ranks. Some of the younger men, especially those in the Colleges of Agriculture and Engineering, are seriously questioning the excellence of their own teaching, and they are coming to educational men for guidance.

Why not? The trouble has been that the college professor has been thinking of school teaching as it was when he was a boy. The truth probably is that greater advances have been made in education in the past decade than have been made in any equal length of time since the savage began to practice initiatory rites. This the college teacher has not known; and he has been unwilling to believe it when it has been brought to his attention. He recognizes progress in other fields, even in those remote from

his own. But in teaching he questions all progress, partly because of his ignorance and partly as a defensive attitude.

Recently, however, we notice a change. The universities are tending to seek as presidents educators rather than chemists, mathematicians, lawyers, or philologists. Of course, a man may be an educator and also any one of these; but the tendency which we are pointing out is the tendency to seek men for presidential positions because of their professional expertness in education.

Again, specialists in education are filling a new type of university position, one in which special administrative ability is required. The position is variously named—provost, dean of administration, even secretary to the president. The duties required, however, are those of educational leadership.

Moreover, the programs of various learned societies are beginning to show the names of prominent educational authorities; and the topics assigned to these men, though elementary as befits the status of their audiences, are nevertheless strictly professional in character. The improvement of college teaching is the idea which seems to actuate the makers of these programs.

Professional education found its earliest application in the elementary field. It has spread to the high school where its battles are almost all won. Its entrance into higher education we believe is assured. We hold this belief not only because the signs of the times point in this direction, but because of the genuine value of the service which education can render. The college teacher is not necessarily a good teacher. Recent visits to college classes have convinced us that some of the poorest teaching is in our higher educational institutions—poorest, that is, from a professional and technical point of view. This condition can be remedied, and some of the college teachers realize this. In bringing about the remedy the educational expert can and should play a part. B. R. B.

Reviews and Abstracts

KELLEY, TRUMAN L. *Statistical Method.* New York, The Macmillan Company, 1923. 390 pp.

We have all waited long and patiently for this book to appear, and now that we have it, there is no doubt left in our minds that Truman Lee Kelley is America's foremost statistician.

The Preface begins as follows:

"This book has been written with a view to serving two needs; that of biologists, economists, educators, and psychologists, who know little of mathematics, possibly care less, and who use statistical methods merely as a device to portray the facts of their group investigations; and that of those in the same fields who resort to mathematics to aid in the discovery of new truths."

"The elementary statistical needs in the four fields mentioned seem to me to be the same and it is my aim to meet those needs and provide a foundation which will serve for advanced work in any one of them."

We doubt if any one who "knows little of higher mathematics" and "cares less" can really use this book conveniently, but for those trained in mathematics it will undoubtedly serve as a most excellent foundation for advanced work.

The book deals with all aspects and concepts of statistical method—central tendency, dispersion, normal distribution, relationship, correlation, partial correlation, and curve fittings,—and unlike preceding books, it furnishes the derivation and proof of all formulas. Here lies its value for the research worker who needs to devise new formulas for the solution of new problems.

The book is written in the style typical of mathematicians. For example, instead of developing the formula for the standard error of the standard deviation of a distribution directly, the author first derives the standard error of any moment, (a very general concept, the moments being Σx, Σx^2, Σx^3, Σx^4, etc.), and then obtains the standard error of a standard deviation (the square root of the second moment), by deduction from the general formula, all of which, of course, is very neat from the point of view of the mathematician.

A free use is made of the Greek alphabet, but the author is thoughtful enough to furnish the Greek alphabet with English equivalents in the appendix and to refer to it in a footnote when he begins to use it.

Appendix A contains eight pages of definitions of the various symbols used in senses special to the field of statistics, but the reader who is not sufficiently grounded in mathematics to know that 5! stands for "factorial 5" and means 1x2x3x4x5 will meet difficulty as early as page 95. Indeed, on this page the author invokes the calculus in the derivation of the equation of the "normal probability curve," and uses it in various places throughout the book.

The typography is excellent. The author is to be congratulated on finding a publisher whose printer could set up the formulas, with the subscripts which are themselves subscripted and the exponents which are themselves exponented. Each formula is numbered; the last is number 326.

Appendix B is a thirteen-page bibliography on statistics and mathematics. Appendix C is a Table of the Normal Probability Integral. The headings of the tables are I, x, z, p, q, etc., but no explanation of their meanings is given in the appendix nor is any reference made there. By consulting the index we find that in Section 27 the table is explained. But to understand Section 27 one must have read and understood Section 26, and to understand Section 26 one must have read and understood Section 25, etc. This is necessarily true in a book of this sort. In other words this is *not* a reference book. One cannot open it up to the topic of the normal distribution or of curve fitting and read comprehendingly. It is a book to begin at the beginning and study, but if the earnest researcher who knows his algebra and has a friend who knows calculus will take the time to study this book thoroughly, he will have laid a foundation for creative work in the field of statistical method.

Note.—For the benefit of those who do not know calculus but have some knowledge of normal distribution and would like to make use of the Kelley-Wood Table, we would state that in it I presumably stands for the word integral, a term used in calculus, and refers to the area of the normal surface of frequency (the figure shown on page 101) between the vertical line at the middle point (mean) and any other vertical line at a distance x from the mean.

Stated in another way, $I =$ the percentage of scores or errors (whatever is distributed), which fall between the mean and some other value represented by x. Thus when $I = .250$, just 25 percent of cases lie between the mean and the value of x under consideration.

The values of x represent the horizontal distances of points on the base line from the mean. The units of x are such that when the distance is equal to σ, (standard deviation of the distribution) $x = 1$. In other words, $1x = 1\sigma$ (in the units of the scores or errors), $2x = 2\sigma$, etc.

The symbol z refers to the height of the vertical line (called the ordinate), at any point on the base line. The value of z in the table corresponding to any value of x shows the probability of a score or error having that value of x relative to the probability of its having some other value of x. That is, if one value of z is twice another, it means that the probability is twice as great that a score or error will fall at the first point in the distribution as that it will fall at the second point.

The first two columns of the table (showing the relation between I, the percent of scores or errors falling between the mean and some other point on the distribution, and x, the distance of that point from the mean in terms of σ) will probably be most frequently used.

ARTHUR S. OTIS

World Book Company

REYNOLDS, ROLLO G. *Newspaper Publicity for the Public Schools.* New York, A. G. Seiler, 1922. 120 pp.

Considerable attention has been given in the past to newspaper publicity for special campaigns. In this monograph particular emphasis is placed upon continuous news service, the purpose of which is to create good will on the part of citizens in order that their support and cooperation with the public school system may be based upon full information.

Doctor Reynolds' investigation included (1) the study of the amount and character of school news found in the daily papers; (2) school news from the editor's viewpoint; (3) school news from the superintendent's viewpoint; (4) school news from the reader's viewpoint; (5) a program for school information service in the daily newspaper.

To furnish the data as to the amount and character of newspaper news, 25 cities were chosen at random representing different sections of the country. In each city the largest daily newspaper was subscribed to for a period of three months. A number count was then made of the newspaper articles dealing with local public schools. Replies were received from 98 editors in cities telling the different kinds of news which met with their approval. School news from the superintendent's viewpoint was determined from the results of a questionnaire which was answered by 250 superintendents representing 45 states. Ten newspaper articles, each having to do with the health of school children, were chosen for the judgment of newspaper readers. Readers were asked to rank them beginning with the most interesting article and ending with the least interesting. Judgments were received from 369 individuals. From these judgments a scale of articles was constructed to show the relative value which the newspaper reader gives to certain school news.

A large number of newspaper clippings were likewise reproduced to illustrate the different types of news which newspapers are willing to print.

In conclusion, Doctor Reynolds gives specific suggestions for an organized school news service and says "A public school publicity or news program should be developed as a part of the whole educational program with the cooperation and advice of newspaper organizations. If it is carefully planned, thoroughly organized, and systematically carried out, it will prove to be of great benefit to the school system. It will be a most powerful agency for securing the interest, cooperation, and support of the public, without which progress is impossible."

P. R. STEVENSON

Ohio State University

WHEAT, HARRY GROVE. ***The Teaching of Reading.*** New York, Ginn and Company, 1923. 346 pp.

This book sets out to do two things important in the training of teachers of reading. First, it assembles the results of scientific investigation, showing to what extent the various problems involved in the reading process may be solved in the light of the evidence in the case. Second, it proposes a course of study which gives to each grade responsibility for certain important reading skills and habits.

In accomplishing the first of these purposes the author quotes from practically everything which has been said about reading for the last twenty-five years, with the idea of showing how much of what has been written is mere opinion as to the nature of reading and how much is based upon reliable data. As one reads the material for the first time, one is rather bewildered by the innumerable quotations. Before the reader is completely lost, however, he comes upon concise summaries which recall to his mind what it is all about.

Soon he develops a technique of reading the summaries first and then reading the previous chapter to find the basis of the conclusions. Teachers of experience will find that the book thus affords a quick way of finding out what has been said and done about the various reading problems.

The second purpose of the book, the establishment of definite goals of instruction for each grade is an attempt that teachers always welcome. As to whether or not stages of progress in reading can be made to correspond with definite school grades is a debatable question. The fact remains, however, that teachers must teach by grades. If, therefore, certain growth periods in reading have been decided upon, each grade must accept responsibility for establishing its appropriate fundamental skills.

In the discussion of these reading skills the use of the term "independent reading" to mean a desultory, cursory sort of reading is rather misleading. It would seem that whether reading is of a recreative type or whether it is of the work type it might be quite independent. Again the distinction between study and reading is a rather arbitrary distinction. The writer says, "reading is not study, reading is getting the meaning easily and rapidly from the printed material and the content must appeal to immediate interests."

However, regardless of such details, which are after all largely questions of terminology, the book furnishes much valuable information to teachers in service as well as to those in training.

ESTALINE WILSON

Assistant Superintendent of Schools, Toledo, Ohio

DICKSON, VIRGIL E. *Mental Tests and the Classroom Teacher.* Yonkers-On-Hudson, New York, World Book Company, 1923. 231 pp.

Doctor Dickson sees the classroom teacher as the chief worker in the educational organization; and the intellectual, social, and physical growth of the child as the chief end to be attained. Therefore, he treats mental tests as a new tool to be used by teachers in increasing their service to children. Nor does he overlook the responsibility of the principal and supervisor; for from beginning to end his book is marked by the quiet, thoughtful, directive attitude of the successful administrator.

Mental Tests and the Classroom Teacher adds no new theory or startling revelation to the literature of mental testing. Its chief asset is its sanity and its proof that an administrator may direct the testing of more than 50,000 children and still not lose sight of the individual child during the years he is in school. Probably no other worker has had the opportunity to test such large numbers of children, and at the same time to work out such a comprehensive following-up program for adjusting the school to their respective educational needs as Doctor Dickson describes in this book.

Moreover, *Mental Tests and the Classroom Teacher* is well written. It is readable. Its author has a philosophy of education and he sees mental testing in its true relation to the whole problem of the education of childhood from kindergarten to the *age of eighteen*. It is not undue praise, to say that it is

the best book yet offered for the superintendent, principal, and teacher to use in studying together the intricate problems involved in the use of mental tests.

J. Cayce Morrison

Ohio State University

Horn, J. L. *The American Elementary School.* New York, The Century Company, 1923. 422 pp.

In writing this volume Professor Horn has undertaken an ambitious piece of work. He has endeavored to bring together something of the development, and even more of the present status, of a large number of problems which are now current in the field of elementary education. He has attempted to do this in order that normal-school students and experienced teachers may get a large, unified vision of this great American institution. His method has been, after an introductory overview of the whole field, to present in successive chapters the various problems which relate to the external relations of the school, to its organization, and to questions which primarily involve instruction. In each chapter he has presented his case in a clear-cut manner, closing with a summary of the points which he has endeavored to make, and following this with a statement of several problems which are very closely allied to the main theme of the chapter and which would furnish material for thoughtful discussion if the book were being used as a text. In addition to this, a selected bibliography is also presented; a bibliography which does not pretend to be complete, but which brings to the attention of the student the material which is most likely to be accessible in a school library.

Not satisfied with this workable organization of the problems under discussion, he closes the book with a chapter in which he states nineteen present challenging problems which school administrators and teachers must face. The most cursory examination of the set-up of these problems in his final chapter will make clear that the author has not only seen his subject in a large way, but has remembered the relatively untrained group for which he is writing.

The book is stimulating; and while for those who have kept fully abreast of the current literature on school administration it contains little that is new, yet for anyone who has any responsibility in the matter of the training of teachers, whether in normal schools or in service, it will be found to be a truly helpful addition to the literature already available.

E. J. Ashbaugh

Ohio State University

Allen, William Sims. *A Study in Latin Prognosis.* New York, Teachers College, Columbia University, 1923. 41 pp. (Teachers College, Columbia University. Contributions to Education, No. 135.)

This study in Latin prognosis is an excellent piece of scientific work which marks a definite advance in a practically virgin field. It is based on experiments with 364 boys beginning the study of Latin. They were given at the outset twenty-one forms of various psychological tests. At the end of the first semester they were given the Briggs series of Latin tests. The correlations

between the two series of tests were then studied. It is to be regretted that no English grammar test was given, for tentative experiments seem to show that such a test has high prognostic value for Latin.

The rather comprehensive Latin test consisted of eleven parts weighted by thirteen Latin teachers. That the test is quite as good as the average examination is shown by the high correlation (.769) between the test scores and teachers' grades.

By a method of multiple correlation the author decides which combination of tests make up the best prognostic series: Briggs Analogies, Thorndike Test of Word Knowledge, and Rogers Interpolation. These give a prediction of .55 to .57. Granting that this combination is the best of the tests which the author used, two questions remain. Is it the best that can be devised? Is it satisfactory? The author would no doubt be the last to answer the first question in the affirmative. He has shown the way for further experiment. The second question will be answered in different ways by different persons. In the first place, Allen himself says that "the prognosis tests are not recommended as a sole guide for admitting or precluding pupils, but as one of several means." In the second place, his combination of prognostic tests predicts success in English and mathematics about as well as in Latin. He says that the predictions for Latin are .588, for English .596, for mathematics .554. But the first figure is based on the Briggs test, not on teachers' marks, as is the case in English and mathematics. For Latin the correlation with teachers' marks is only .485. This is quite a bit lower. But even if all the coefficients of correlation were about the same, serious doubts as to the value of the prognostic test combination are raised by the fact that it predicts equally well for subjects as diverse as those mentioned. The test combination is really suitable not for prognosis in any one subject but rather for prognosis in a group of academic subjects. It may serve, with other data, in advising pupils for or against an academic course, but hardly for or against Latin any more than English or mathematics.

B. L. Ullman

State University of Iowa

Rusch, Louise C. *The Socialized Recitation in English.* Chicago, The Plymouth Press, 1923. 88 pp.

A pure description of the socialized recitation never brings its real spirit to the teacher who is unfamiliar with the procedure. Visiting some teacher who is putting socialized recitations into practice is the best way to catch the finer side of the work, the spirit which makes it different from the usual type of teacher-led class. The difference lies primarily in the movement, the exhilaration, the life, which the pupils put into the work that they, themselves, have planned and are carrying on.

Unfortunately, teachers seldom have an opportunity to observe the teaching of others, and almost never to observe desirable or unusual teaching. Miss Rusch has used the next thing, the verbatim report. Such a reproduction, while lacking some of the keen personal element, the flash of minds in the interchange of ideas, the magnetism of voice and eye, has an advantage in that the

teacher or student may review the actual words used and find a new source of information and of ideas that would be impossible otherwise. She will have, too, a permanent form for comparison.

The author of this little book presents the verbatim reports of lessons that actually took place with real children "of not above average ability." The reports show what actually can be done, for it has been done. While the lessons are English lessons, the suggestions are equally applicable to other subjects. However, the real contribution of the author is the collection of reports illustrating a method of recitation.

DWIGHT H. STEVENSON

Ohio State University

SYMONDS, PERCIVAL MALLON. *Special Disability in Algebra.* New York, Teachers College, Columbia University, 1923. 188 pp. (Teachers College, Columbia University. Contributions to Education, No. 132.)

Upon the determination of individual differences in various school subjects a great deal depends—matters of guidance, curriculum offerings, the grouping of pupils, promotion schemes, and many other problems of educational theory and practice. Much has been written on this subject, but until the last decade little light has been shed upon the extent to which capacities differ in school subjects. Indeed, the reliable studies of this sort may still be counted on one's fingers. The problem needs to be investigated with reference to each school subject.

One of the most careful studies of disabilities in a school subject is reported by Doctor Symonds in his *Special Disability in Algebra.* The author defines special disability as "discrepancy between ability in some trait and intelligence when compared with the average performance of individuals as measured by some test." Twenty-two pupils who had had trouble with algebra during the previous year were given the following tests: Army Alpha, Woody-McCall Mixed Fundamentals, Hotz Algebra, Rogers' Tests of Mathematical Ability, and the Thorndike-McCall Reading Scale.

Two methods of approach were employed: (1) The study of individual cases by means of psychographs and the results of a questionnaire; and (2) the calculation and use of certain derivations from correlations properly corrected, to measure the amount of special disability.

The latter method exhibited for each pupil the amount of special disability in terms of the standard deviation of Alpha scores. This was determined by first finding the amount by which the individual's Alpha score exceeded the mean score for the whole population. From this amount was subtracted the superiority of the individual as it would result if the measure of his intelligence was taken as corresponding to his performance on the algebra scale. In this subtraction both minuend and subtrahend were expressed in terms of the standard deviation of Alpha scores for the whole population.

Measurable special disability in algebra was found to exist for at least twenty of the twenty-two cases studied. It ranged from 0.45σ to 3.39σ.

In addition to these facts, Doctor Symonds presents at the close of his discussion some practical suggestions on such points as the course of study, textbooks, class assignments, class periods, and the question of who should study algebra.

Incidentally to the main problem, the author makes a contribution to statistical procedure as ordinarily followed. He shows the necessity for the correction of coefficients of correlation, and clearly outlines the method of making such corrections on the basis of the reliability and selection of data.

While the study is by no means the last word to be said concerning special disability in algebra, the results are truly significant; and the methods and formulas employed—for the full and clear description of which the author should be especially commended—will be very suggestive to others who may make similar studies.

H. R. Douglas

Stanford University

News Items and Communications

This department will contain news items regarding research workers and their activities. It will also serve as a clearing house for more formal communications on similar topics, preferably of not more than five hundred words. These communications will be printed over the signatures of the authors. Address all correspondence concerning this department to Doctor E. J. Ashbaugh, Ohio State University, Columbus, Ohio.

Under the new laws governing education in Austria adopted since 1918, all members of school boards must themselves have been teachers. We wonder in what direction schools would be influenced in this country if a similar law were in effect.

According to Nebraska's new curriculum which will go into effect in September, 1924 small rural schools will reorganize the eight elementary grades into four "form schools" of two grades each. These groups will take up the different grades of work in successive years.

Doctor F. J. Kelly, for some time professor of Administrative Research and assistant to the Chancellor at the University of Kansas, has been made Dean of Administration, Assistant to the President, and Director of the Summer Session at the University of Minnesota.

The Alaska School Bulletin issued by the Territorial Department of Education announces three elementary-school contests for the current year. The contests are in spelling, rapid calculation involving whole numbers, and silent

reading. The contests are open to pupils in the sixth, seventh, and eighth grades.

The Kansas State Normal School at Emporia which conducted such a successful scholarship contest last spring has announced its second contest for May 3, 1924. A news item on the nature of the contest can be found in the October issue of the *Journal* on page 273.

Doctor Roy L. Davis, one of the pioneers in the field of visual instruction, is now associated with the Society for Visual Instruction, Chicago, Illinois. He has had a wide experience in the theoretical, as well as the practical use of visual aids in the classroom. Under his supervision the Society is developing a comprehensive system of visual instruction that will be of vital interest to the school public.

We have just received from Professor Alexander Inglis of Harvard University a series of Latin tests and an English vocabulary test. The Latin test consists of general vocabulary, syntax, and morphology. Each of these is available in five different forms. The Latin vocabulary and English vocabulary tests have been well standardized, and tentative norms are in process of derivation for the others. The tests may be secured from the Harvard University Press, Cambridge, Massachusetts.

The Board of Education of Newark, New Jersey, has planned to remove some of the ultra modern policies now in vogue. The first steps apparently are to eliminate all special teachers in junior high-school grades, and to reorganize the high schools in such a way as to eliminate duplication in teaching. It has already been decided not to appropriate funds in the future for the purchase of uniforms, instruments, and equipment for bands and fife and drum corps.

New Jersey requires candidates for entrance to normal schools to take an examination in each of the following subjects, spelling, English, and arithmetic. The results with 1,215 candidates in June, 1923 show as follows: 95.7 percent passed spelling; 82.8 passed English; and 79.0 percent passed arithmetic. Two and one-half percent of the candidates failed in all three subjects, and 31½ percent failed in one or more subjects, thus failing to secure entrance.

The J. B. Lippincott Company announces the publication of a new speed of reading test by Chapman and Cook. The tests consist of thirty paragraphs in each of which there is one word that spoils the meaning. The task of the child is to read the material as rapidly as possible, find this irrelevant word, and cross it out. The tests were designed for grades four to eight, inclusive, and norms have been determined not only in terms of median but of four gradations in each direction from this median.

The County Board of Education, Mobile, Alabama, is publishing an eight-page paper through the office of the County Superintendent. The cost of the paper is largely supported by advertisements, and the content consists of desirable education information which the school authorities wish to disseminate to the school people of the county. A large number of letters from children of various grades in the different schools present evidence of the work that is going on, and the needs of the schools.

The Commonwealth Fund has appropriated a sum for support of the study of history textbooks. The particular texts to which consideration will be given are those which are written for use in the sixth grade, and which have an "Old World" background. Dr. Raymond Franzen of the University of California will be in charge of the project. The texts which he will examine will be evaluated primarily in terms of the criteria of interest, comprehension, and permanent value of subject matter.

We note in the October issue of the *Alabama School Journal* the report on adult schools for 1922-23, which shows that under the direction of the Division of Special Education 5,359 pupils (adults) were taught by 202 teachers in 178 schools in 22 counties of the State. One county enrolled 392 in 11 schools. The pupils learned citizenship, sanitation, and recreation, as well as the three R's. During the term of sixty hours a large portion of those who had never learned to write mastered the art sufficiently well to write a letter.

Philadelphia school authorities are using reinforced concrete construction for school buildings instead of brick and steel and are finding that this occasions a very substantial saving of both time and money. Since there are few of our larger cities which are not actually two or more years behind in their building program, and since concrete buildings can be built in less time than brick and steel, the use of the former material is very important in any attempt to catch up with building needs.

The city of Detroit is using a series of reading tests devised by a number of teachers working under the guidance of Miss Claudia M. Parker of the reading department. These tests are designed for use in grades II-A to IX-A, inclusive. Each test consists of 12 paragraphs upon which two questions are asked. One question is purely factual and the other is interpretative. In all cases the multiple-choice type of response is presented, and the child needs only to underscore the correct word. The test is a time test, and the scoring is in terms of the number of questions attempted and the percent of accuracy.

The last legislature in California passed a law which required that the constitution shall be taught in all schools and grades above the seventh. It placed in the hands of the State Department the responsibility for deter-

mining the time allotment and the nature of the work. This has been done as follows: grade VIII, two periods weekly; grades IX and X, citizenship developed as a part of the regular English work; grades XI or XII, 72 periods during the year devoted to United States history; grade XII, 72 periods to be given to a study of the constitution and the working of the government.

St. Cloud, Minnesota, is continuing its public evening school which meets twice each week for sixteen weeks. Eight courses are offered in trades, seven in home making, two in citizenship and beginning English, five in commercial, and seven in academic work. A small enrollment fee is charged, but this is returned if the student is in attendance seventy-five percent or more of the class sessions. The interest of a large number of St. Cloud's adult population in the evening-school program is evidenced by the fact that many have attended for three and four years.

The following facts concerning the high schools of New Jersey will be of interest to any city, county, or state superintendent who would like to make comparisons with the situation in their own territory. Twenty-four percent of the students in the New Jersey high schools are taking classical courses; 30 percent take commercial courses; and 18 percent, modern languages. Only five percent are taking anything that can be classified as strictly vocational subjects. But three-fourths of one percent are pursuing a course especially planned to prepare for normal school. Twice as many girls as boys are taking commercial curricula, and practically one-third of the cost of high school is devoted to preparation of young people for a business career.

The Detroit school system has become well known throughout the country for the numerous problems which it has under investigation. Among the more recent ones in which school people everywhere are interested is the problem of the relative efficiency of part-time and whole-time sessions in the first four grades. The superintendent has appointed a committee of district principals who will organize a number of controlled investigations in which the single variable will be the whole- or part-time session. Among the factors which they are endeavoring to stabilize in the investigation are: Experience, tenure, and rating of the teacher, and the language and social status of the children in the various rooms.

T. L. Torgerson, director of educational measurements in the public schools of West Allis, Wisconsin, has furnished us with copies of an individual standard test record card and a class progress sheet. The test record card is intended to be cumulative throughout the child's career, and contains space for achievement tests in arithmetic, reading, spelling, language, geography, and history. There is also space for a record of intelligence tests. The class progress sheet is intended to be posted in the classroom and will show in graphic manner the status of each individual child in the room with respect to grade standards.

The well-informed *New York Evening Mail* announces that in order to take care of the high-school pupils in New York City, the Board of Education will have to erect fourteen new buildings. This statement, it is reported, is made on the authority of Doctor John E. Ferguson, chairman of the Board's Committee on Sites and Buildings. In view of the fact that it costs approximately $3,000,000 to erect the type of high school which the board has adopted, the astonishing cost of a program which will do no more than wipe out part time may be appreciated. This has no reference to the buildings or sittings needed to do away with double sessions, or to take care of natural increase.

Dean John W. Withers of New York University, in a recent address before the students in the Maxwell Training School for Teachers, New York City, emphasized a phase of our unequal educational opportunity that is often overlooked. Attention has frequently been drawn to the fact that the country child is receiving fewer days' instruction at the hands of more inefficient teachers than the child in the city. But Dean Withers points out that this is not a condition dangerous to the rural communities alone but is also an element of danger to the cities because large numbers of these children will undoubtedly spend their adult life in the city. Thus these children of today who are receiving inadequate educational training in the country will in a few years help to determine the policies of our urban communities.

This department is not much given to reviews of articles in current journals, but the one by Professor J. W. Searson of the University of Nebraska on "Public Contacts of the English Teacher" in the September number of the *English Journal* is of such significance that the temptation to call attention to it is too great to resist. English teachers, high-school principals, and superintendents of schools will do well to read and ponder over the article. When news items written by pupils without help show a larger percent of items up to the standard of other news in newspapers than do those written by English teachers, by high-school pupils under teacher direction, and by school superintendents, a serious analysis of the English teaching in this field needs to be made.

Dr. Thomas E. Finnigan was, as most people think, subjected to a "raw deal" by Governor Pinchot last spring. As a condition of his reappointment to the position of State Commissioner of Education, the Governor required him to place in his hands an undated resignation. This Doctor Finnigan refused to do; and his refusal resulted in his termination of service. It is with some satisfaction, therefore, that we learn of his association with the National Transportation Institute at a salary which, according to an announcement of Secretary Crabtree of the N. E. A., amounts to $18,000 a year. In his new position Doctor Finnigan, it is said, will "organize a program of education to enlighten the people on the problems of transportation and distribution."

We are in receipt of the Public Survey of Oconee County, South Carolina, made by the Division of Education of Clemson Agricultural College in coopera-

tion with the County Board of Education of that county. The survey contains three hundred pages, and is arranged under the following heads: organization; teachers; enrollment, attendance, and progress; classification and promotion; instruction in primary grades; the school and community; buildings, grounds, and equipment; and supporting the schools. The monograph is well written, nicely printed, and well bound. A great amount of data is presented in tabular form, but each chapter closes with a summary and a list of definite recommendations. It is well illustrated with graphs and plates which portray vividly the conditions found. On the whole, it is one of the very best county surveys which has ever come to our attention.

We have received from H. D. Rinsland, Department of Research and Guidance, Ardmore, Oklahoma, a copy of an individual test record card which he is using. The card provides, besides the name, age, and grade location of the child, a space for recording scores for three tests in reading, two in arithmetic, one each in nature study and science, history and literature, language and dictation. These scores are to be translated into educational age. Space is provided for educational rank, pedagogical rank, pedagogical age, mental age, intelligence quotient, educational achievement, educational ratio, and promotion achievement ratio. There is a place to indicate the group assignment of the child into X, A, B, C, and Y classes, X indicating a very retarded group segregated for special work, and Y, the very bright group at the upper range of the distribution.

[Will others who have devised cards for test score records please send me a copy?—*Editor*]

Mr. W. H. Hughes, director of Educational Research for the Pasadena public schools has organized what he calls an "Independent Seminar." This group consists of some of the best and most capable school people of the city. The problem on which the seminar is working is: (1) to determine as scientifically as possible the traits and attitudes that the public schools are interested in developing and (2) to take up individually an intensive study of a single trait or attitude in the list which has been developed. Each member will study this trait or attitude with a view to ascertaining its objective evidences as expressed in behavior. His chief concern, however, will be to find out the methods and types of organization which are most favorable for its development in the school. It is expected that a manual for the teachers of Pasadena will grow out of the work of this seminar.

Doctor David Snedden of Teachers College in an address recently delivered before the New York Society for the Experimental Study of Education discussed the general question of what to teach in civics. He stated that civic education has a major purpose in schools, and called attention to the following lines of approach: (a) formal or didactic civics, (b) American history, (c) historical and other patriotic reading, (d) social life, (e) current events, (f) concrete economics, sociology, etc., (g) dramatic projects, (h) service

projects, and (i) social-science problems. He called attention to the need of analyzing the objectives into workable units and then the elimination through classification into those which are indispensable, very important, important, not very important, and negligible or worthless.

From the *Kansas City School Bulletin* of October we learn that an arrangement had been made between the Board of Education and the churches of that city by which the school shares the time of the children with the churches which wish to give religious instruction. Children are excused from the public schools for one-fourth day each week. First and second grades attend the church school during the first quarter of the day; third and fourth, during the second quarter; fifth and sixth, during the third quarter; and seventh and eighth, during the fourth quarter.

Attendance at religious classes is not compulsory, but a permit to the child is given upon request of the parent or guardian. The churches provide competent teachers, supervisors, equipment, and courses of study and admit those who wish to attend regardless of church affiliation. Courses in ethics and citizenship are provided by the school for those children who are not attending the church school.

Schools, an educational publication of Teachers College, Sidney, Australia, brings two items of interest. The first is a comparison of the average ages of pupils in the first six grades of the elementary school from 1914 to 1921. This comparison shows a steady decrease in the ages of pupils in each grade from year to year. The average decrease for all grades is slightly more than one and one-half years. Apparently children are passing through the grades more rapidly than they were in 1914, because, while the average age of the first grade has decreased less than a year, that of the sixth has decreased a full one and one-half years.

The other article of interest is "An Analysis of Types of Error in the Fundamentals of Arithmetic." The article is continued from the previous issue. It shows us that this type of study, which is quite prevalent with us, is also holding the attention of educators in Australia. A comparison of the findings of the writer with those of investigators in this country shows more similiarity than difference.

W. V. Bingham, of Carnegie Institute of Technology, has been elected Editor and L. L. Thurstone, of the Bureau of Public Personnel Administration, Associate Editor of the *Journal of Personnel Research*. C. S. Yoakum continues as managing editor of this magazine which is now in its second volume. This *Journal* is devoted to the scientific study of personnel. It publishes original researches in the applied sciences that contribute to our knowledge and our effective direction of people at work. It is the official organ of the Personnel Research Federation whose purpose is the furtherance of research activities pertaining to personnel in industry, commerce, education, and government wherever such researches are conducted in the spirit and with the methods

of science. It brings together workers in the biological and medical sciences, psychologists, psychiatrists, engineers, and economists who are making significant contributions to more exact knowledge of human beings at work and of methods of making effective and satisfying adjustments between men and their work.

It is not without significance that the pamphlet issued by the National Education Association in connection with the observance of Education Week was a research bulletin. Not long ago, the department in calling attention to the observance of such an occasion would have broad-casted a lot of fine phrasing whose merit would have been its literary style and whose defect would have been a lack of substance. Nothing of this sort occurs in the case of the research bulletin to which we refer. Five questions are suggested in connection with the observance of Education Week. On each of these questions important fact material is given. The questions are: (1) What are the weak spots in our public school system? (2) What national defects result from the weak spots in our public school system? (3) How may our public school system be strengthened? (4) Can the nation afford an adequate school system? (5) Do good schools pay?

We hope that much of the thinking of our citizenry and not a little of the public speaking which took place in connection with Education Week was vitalized by the research bulletin which the N. E. A. has issued.

The University of Kansas has initiated the Kansas Studies in Education, Volume I, Number 1 with the Shawnee-Mission Rural High School by F. P. O'Brien and T. J. Smart. The bulletin reports a survey made upon request of representatives of school boards of fourteen school districts in Johnson County, Kansas. A situation had arisen in which the high-school facilities of these districts could not be provided, as formerly, due to the increase in costs and tax limitations. The writers call attention to the fact that assistance was asked because of financial pressure rather than from any dissatisfaction with instruction offered their children in a "D grade" high school. An analysis of the districts as to population, school attendance, valuation, occupations of parents, and vocational choices of children were the data upon which a proposed organization of a new district was made.

The survey was made in 1920-21; and at the time the report was ready for publication, the recommendations had been followed to the extent that a new district of large size had been created, a twelve-acre site purchased, a modern rural high-school building constructed, and a modern school actually set in operation. Whereas, at the time the survey was made the school levy in the districts which were later consolidated varied from .75 of a mill to 25 mills, the new district is now maintaining a high type school on a levy of 3.6 mills.

Intelligence Tests in the Civil Service

One of the newest departures in the field of current literature is the publication of the Bureau of Public Personnel Administration, which comes

out under the title "Public Personnel Studies." Volume I, Number I, dated October 15, 1923, has recently come into the hands of the writer. The function of the publication is best explained by a statement concerning the organization which is responsible for it.

The Bureau of Public Personnel Administration is an organization established through the efforts of civil service administrators to serve civil-service commissions. It is supported independently and is not officially connected with any federal, state, county, municipal, or other governing body. It purposes to serve as a clearing house for information relating to personnel administration in the public service; to develop methods of personnel administration through original investigations and experiments; and to publish the results of such studies in the most effective and expedient manner. The Bureau is under the direction of W. F. Willoughby who is assisted by a staff composed of Fred Tilford, L. L. Thurstone, Lewis Merriam, Thelma Gwinn, and L. J. O'Rourke.

The studies of which this number is the first are to be published "at frequent though somewhat irregular intervals." The general policy will be "to publish only such conclusions as are supported by good current practice or other facts or by sound opinion." For the present these pamphlets will be, as this one is, in mimeographed form with a printed cover.

Doctor Thurstone has the honor of being the author of the first of the series, this number having the title "Intelligence Tests in the Civil Service." In it he has given a very simple but discriminating discussion of the construction and function of the various types of mental tests. The treatment is so simple that the wayfaring man, though a fool, cannot err in getting his meaning.

Although the article is addressed to civil-service administrators, school people can truly profit by a careful reading of the article. The best of it is that Doctor Thurstone has promised a further study in the immediate future, dealing with a comparative study of intelligence tests now in the field. He is also proposing to conduct experiments to determine the weight to be given to intelligence tests and to other tests for particular positions. We shall look forward with keen anticipation to the publication of these studies.

DWIGHT H. STEVENSON

Ohio State University

Plotting Graphic Distribution Charts With a Typewriter

Figures 1 and 2 illustrate a method of making graphic distributions of test scores, school marks, etc., which has worked well in our office. To statistical students the advantages will be apparent at a glance; the important point, however, is that teachers untrained in statistical methods quickly learn to read a single standardized form. In operation it is absurdly simple, yet it makes use of scientific principles. No claim is made to originality unless it be in outlining a definite way of making use of well-known methods in a form so designed that final copy can easily be made on an ordinary typewriter.

A form is first mimeographed on standard 8½ x 11 mimeograph paper. This form is that shown in Figure 1 without the figures and check marks

which are obviously handwork. A supply of these forms is always on hand. To distribute scores take a blank form and make in pencil the entries shown in Figure 1. This is then copied by the typist to appear as shown in Figure 2. The slightly technical points involved may be explained as follows:

STEP 1: CHECKING OFF THE SCORES— THE GRAPHIC DISTRIBUTION

(a) Determine approximately the range of the scores and group them sufficiently so that they will be included within the range of the form (*See* Column 1, Figure 1). The same grouping should be adhered to when distributing the same test for other groups.

Figure 1

GRAPHIC DISTRIBUTION CHART NO. 7

INSTITUTION Public Schools #32, 34 TESTED December 1922.
GROUP 7B Classes TAKEN BY Dorian Richards
MEASURES IN Todays Clerical Test.

NUMERICAL DISTRIBUTION				GRAPHIC DISTRIBUTION (Each check mark represents the position of 1 pupils) 5 10 15 20 25
Percentile Rank	Sums	Freq.	Score	
(4)	(3)	(2)	(1)	
.002	0	0	4	
.002	1	1	6	v
.006	3	2	8	vv
.011	5	2	10	vv
.019	9	4	12	vvvv
.028	13	4	14	vvvv
.041	19	6	16	vvvvvv
.054	25	6	18	vvvvvv
.073	34	9	20	vvvvvvvvv
.095	44	10	22	vvvvvvvvvv
.117	54	10	24	vvvvvvvvvv
.140	65	11	26	vvvvvvvvvvv
.168	78	13	28	vvvvvvvvvvvvv
.207	96	18	30	vvvvvvvvvvvvvvvvvv
.255	118	22	32	vvvvvvvvvvvvvvvvvvvvvv — 25%
.319	148	30	34	vvvvvvvvvvvvvvvvvvvvvvvvvvvvvv
.387	179	31	36	vvvvvvvvvvvvvvvvvvvvvvvvvvvvvvv
.462	214	35	38	vvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvv — 50%
.544	252	38	40	vvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvv
.618	286	34	42	vvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvv
.687	318	32	44	vvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvv
.756	349	31	46	vvvvvvvvvvvvvvvvvvvvvvvvvvvvvvv
.814	377	28	48	vvvvvvvvvvvvvvvvvvvvvvvvvvvv — 75%
.862	399	22	50	vvvvvvvvvvvvvvvvvvvvvv
.898	416	17	52	vvvvvvvvvvvvvvvvv
.927	429	13	54	vvvvvvvvvvvvv
.946	438	9	56	vvvvvvvvv
.961	445	7	58	vvvvvvv
.972	450	5	60	vvvvv
.981	454	4	62	vvvv
.987	457	3	64	vvv
.992	459	2	66	vv
.994	460	1	68	v
.996	461	1	70	v
	461		72	
.998	462	1	74	v
	462		76	
1.000	463	1	78	v
			80	
			82	

Total Number 463 Range: (See Freq. Column Above)
25%ile 33.8 Median 40.9 Q 7.0 75%ile 47.9

(b) Having determined this point, the next step is merely to make a check mark in the proper row under "Graphic Distribution" for each pupil's test score. (*See* Figure 1). Two or more checks may be entered in each square. If the group is large and these checks "overflow" at the right margin of the graph, cut a second sheet of blank form and attach as an extension.

STEP 2: EXPRESSING THE SAME FACTS NUMERICALLY—THE NUMERICAL DISTRIBUTION

After the pencil graph is completed the steps are as follows:

(a) Fill in the *Frequency* column by counting the check marks in each row and entering same under *Freq.*

(b) Fill in the *Sums* column which is made up of the cumulative totals beginning at the top. This shows at a glance the total number of pupils below

Figure 2

GRAPHIC DISTRIBUTION CHART NO. 7

INSTITUTION: Public Schools # 32, #34 TESTED: December, 1922.
GROUP: 7 B Classes EXAMINER: Bureau of Research
MEASURED IN: "Iowa" Clerical Test

Percentile Rank	Sums	Freq.	Score	GRAPHIC DISTRIBUTION (Each star represents the position of 1 pupil)
.00	0	0	4	
.002	1	1	6	*
.006	3	2	8	**
.011	6	3	10	**
.019	9	4	12	****
.028	13	4	14	****
.041	19	6	16	******
.054	25	6	18	******
.073	34	9	20	*********
.096	44	10	22	**********
.117	54	10	24	**********
.140	65	11	26	***********
.168	78	13	28	*************
.207	96	18	30	******************
.255	118	22	32	********************** ← 25%ile
.296	148	30	34	******************************
.357	179	31	36	*******************************
.462	214	35	38	*********************************** ← Median or 50%ile
.544	252	38	40	**************************************
.618	286	34	42	**********************************
.687	318	32	44	********************************
.754	349	31	46	******************************* ← 75%ile
.814	377	28	48	****************************
.862	399	22	50	**********************
.898	416	17	52	*****************
.927	429	13	54	*************
.946	438	9	56	*********
.961	445	7	58	*******
.972	450	5	60	*****
.981	454	4	62	****
.987	457	3	64	***
.992	459	2	66	**
.994	460	1	68	*
.996	461	1	70	*
	461		72	
.998	462	1	74	*
	462		76	
1.000	463	1	78	*

Total Number 463 Median 40.9 Range: (See Freq. Column Above)
25%ile 33.8 Q 7.0 75%ile 47.9

any given score. By subtracting the entry in the *Sums* column opposite any score from the last entry in this column it likewise shows the total number exceeding that score.

(c) Fill in the *Percentile Rank* column by first calculating the reciprocal of the total number of cases. Beginning at the top add for each row as many times the reciprocal as there are frequencies for that row; or stated differently —for each row multiply the entry in the *Sums* column by the reciprocal and

enter the product in the *Percentile Rank* column. If an adding machine is available the calculation of percentile rank is made extremely simple by the method described by Thurstone.[1]

Summary at Foot of Sheet.—For the summary at the bottom of the sheet it is now merely necessary to copy the items wanted. Thus the approximate 25-percentile, the 50-percentile (median) and the 75-percentile can be read directly from the *Percentile* column. Any other percentiles may, of course, likewise be noted and copied at the bottom if desired, e.g., the 10-percentile and the 90-percentile. The total number of cases is the last number in the *Sums* column. The range is apparent in the frequency column. "Q" is immediately available by subtracting the 25-percentile from the 75-percentile and dividing by 2.

This is the form we have been using for ordinary rough calculations. In case it is desired to obtain the arithmetic mean and the standard deviation, it is only necessary to add two more columns at the left of the form here presented—the first one a "Deviations" column and the second one a "Deviations Squared" column.

STEP 3: FINAL COPY

Since the blank conforms to standard typewriting spacing, a clear copy of every entry as shown in Figure 1 can next be made by any typist, and the net result will be as shown in Figure 2, in which all column and row lines are omitted. This method immediately makes available all the standard typewriter duplicating methods for the final copy. Carbon or mimeographed copies of these distributions can thus be easily produced. Furthermore, all ordinary distributions thus appear on standard 8½ x 11 sheets, which is a great convenience in filing.

Note to Typist.—If the longest row of check marks (*See* Figure 1) extends beyond the right hand margin of the form, it is only necessary to group the entries sufficiently so that they will fall within the space available. That is, let one star stand for *more* than one pupil, e.g., two pupils, three pupils, four pupils or five pupils, uniformly. Do not group more than necessary. Any number of cases can thus be entered on standard 8½ x 11 paper.

GENERAL ADVANTAGES

The chief advantages are probably obvious but may be summarized as follows:

By making use of a standardized form, labor is reduced and the process simplified. By adopting typewriter methods, graphic distributions which ordinarily are technical and involve mechanical drawing can be produced by a typist. Most important of all, the practice of reporting entire distributions instead of mere averages is encouraged. In Figure 2 we have the entire story, a numerical and graphic presentation of *all* the facts, yet they are included on a single page. Like any other distribution it tells at a glance the highest

[1] Thurstone, L. L. "The Calculation and Interpretation of Percentile Ranks". *Journal of Educational Research*. 6:227, October, 1922.

and lowest scores, the typical scores, the typical deviation from the median, the number of individuals receiving each score, the number exceeding any given score, the number below any given score, etc. Such a record conveys an infinitely better notion of the facts in question than can possibly be had from merely quoting averages. Furthermore, the method of building up a surface of frequency by actual units is more meaningful to the untrained person than merely describing the outline of the curve.

Such a record as the one described tends to convey even to the untrained something of the meaning of a variable.

There is a further point of deeper significance involved—namely, that the practice of keeping separate the distributions of groups of different descriptions is encouraged when distribution charts are easily made. The distributions of *various* well-defined groups are highly desirable for adequately interpreting any individual's score. Chapman has ably pointed out the shortcomings of norms derived from mixtures of many sorts of populations. Files of such distributions for well-defined groups become significant as special norms for reference. A pupil's test score becomes far more meaningful when it can be identified as equal to a definite position in the graphic distributions of specific groups.

JOHN L. STENQUIST

Director of Research
Public Schools, Baltimore, Md.

NEW PUBLICATIONS

Schutte Scale for Rating Teachers

A scale for use in making a thorough and detailed rating of a teacher. Five general traits are rated: Personal and social qualities, Cooperative qualities, Leadership, Scientific and professional attitude, and Teaching ability.

The scale consists of 86 questions grouped under the five heads and used as a basis for rating. The following are samples of the questions:

Is she careful in dress and social niceties?
Is she tactful in dealing with pupils and patrons?
Does she have a spirit of cooperation and team work?
Does she show an interest in professional reading?

The booklet gives complete directions for making the ratings and for showing the results graphically as a profile.

The scale may be used to excellent advantage by passing one copy to each teacher and asking her to rate herself. This has a most wholesome effect upon the teacher and yields information of value to the principal.

Price per package of 25: $1.00 net.

Morrison-McCall Spelling Scale

A 16-page booklet containing lists of words for use by the teacher in testing spelling. There are 8 equivalent lists of 50 words each which enable the teacher to test the spelling ability of a single class 8 times with different words and thus measure progress.

The words chosen are found in the Ayres Spelling Scale and in the 5,000 commonest words of Thorndike's Word Book. They are arranged in order of difficulty in each list. A sentence is given to illustrate each word in order to avoid misunderstanding.

The booklet contains complete directions for giving the tests, scoring the papers, recording results, and finding the spelling age of the pupils. The norms used are based on the testing of 57,000 pupils.

Price per booklet: 25 cents net.

WORLD BOOK COMPANY

Yonkers-on-Hudson, New York **2126 Prairie Avenue, Chicago**

AMERICAN BOOK COMPANY

EDUCATIONAL PUBLISHERS

NEW YORK, 100 Washington Square
CINCINNATI, 300 Pike Street
CHICAGO, 330 East 22nd Street
BOSTON, 63 Summer Street
ATLANTA, 2 North Forsyth Street

Address all communications to the nearest office of the company. Descriptive circulars, catalogues, and price lists will be mailed on request to any address.

Your correspondence is solicited and will have prompt attention.

Whatever Your Question

Be it the pronunciation of **Bolsheviki or soviet,** the spelling of a puzzling word—the meaning of **broadcast, sippio, agrimotor, Air Council, Ruthene, capital ship, Esthonia, rotogravure, hot pursuit, Fascista, aerial cascade, Blue Cross,** etc., this **Supreme Authority**—

WEBSTER'S NEW INTERNATIONAL DICTIONARY

contains an accurate, final answer. 407,000 Words. 2700 Pages. 6000 Illustrations. Regular and India-Paper Editions.

G. & C. Merriam Co., Springfield, Mass.

Write for specimen pages, prices, etc., and FREE Pocket Maps if you name this publication.

Title: **Smaller Classes or Larger**
(Journal of Educational Research Monographs No. 4.
B. R. Buckingham, Editor.)
127 pages, 6x9. Cloth. $1.25 postpaid.

Author: **P. R. Stevenson**

Subject: **An extensive study of instruction to determine the relation of class-size to the efficiency of teaching. In certain subjects the size of class does not decrease efficiency of instruction.**

Superintendents and School Directors will find this of value.

PUBLIC SCHOOL PUBLISHING COMPANY
Bloomington, Ill.

The Hudelson Typical Composition Ability Scale

This Composition Ability Scale puts into the hands of teacher, pupil, and superintendent a means of learning what is the teaching situation and also of guidance and stimulation toward better results. (a) It enables the writing proficiency of each pupil and of classes and schools to be objectively measured, and (b) accurately compared. (c) The pupils may then be fairly classified into groups of like abilities and (d) justly graded. Such a scale of graded compositions (e) provides pupils an incentive and immediate goal for self-competition. (f) The supervisor is enabled to test impartially the various methods of teaching and the results of each; and (g) the teachers are helped to discover and to train their ability in judging general merit.

The composition selected for this scale is an adaptation from Thomas B. Aldrich's *Story of a Bad Boy*. It was chosen through an investigation of the results from thirty-two representative assignments and the Typical Composition Ability Scale constructed. This scale is printed on a strong sheet of very convenient size to use on the desks of pupils. Every pupil should have a copy. The teacher requires in addition a copy of the *Teacher's Handbook* of sixteen pages which gives full instructions, standards, assignments by which to determine progress and practice exercises.

Hudelson Typical Composition Ability Scale, package of 25—$1.00.

No. of copies

Price single copy—10c.....
Teacher's Handbook, price per copy—10c.....
Sample set, by mail 20c.....

Order from the

PUBLIC SCHOOL PUBLISHING CO.
BLOOMINGTON, ILLINOIS

RICH'S CHEMISTRY TEST

This series of tests is most carefully prepared in both material covering the subject and in arrangement for giving under fixed conditions and for objective scoring. T scores are used and are strictly comparable from one test to another. No computations are required as T score tables are provided for each test, giving directly the T score for each number of correct answers. The attainment of any one pupil is numerically comparable with that of others, and classes as a whole or in groups may be compared. Attainment is measured in (a) ability to think; (b) information; (c) ability to solve numerical problems; and (d) habits and knowledges acquired from work in the laboratory. The tests are arranged to call out these abilities in each cycle of six questions. Tables are given, also, for converting T-scores into the usual percent grading or grading by letters.

Standards are available for attainment from the middle of the first semester of study through four semesters of work. Two tests, called Gamma and Epsilon, are now on the market. If desired, the two can be used on alternate desks in testing the same class to prevent "cribbing." The manual for the teacher is required but should be ordered separately, as the one manual is used with both of the present tests. The tests are furnished in packages of 25 tests.

Rich's Chemistry test Gamma, price per package of 25—$1.00.....
Rich's Chemistry test Epsilon, price per package of 25—$1.00.....
Rich's Chemistry test Teacher's Manual price 15c each.....
Sample set 20c.....

Order from the

PUBLIC SCHOOL PUBLISHING CO.
BLOOMINGTON, ILLINOIS

OHIO STATE UNIVERSITY
COLLEGE OF EDUCATION

SPECIAL CURRICULA OFFERED DURING 1923-24

Fine Arts
Home Economics
Industrial Education
Elementary and Rural Education
Commercial Education
Education-Science Nursing
(For school nurses and teachers in schools of nursing)
Vocational Education
Physical Education

General curricula in all the standard academic subjects are also offered.

All curricula of the College of Education require not less than four years for completion and lead to the degree Bachelor of Science in Education.

Graduate work leading to the degrees of Master of Arts and Doctor of Philosophy is offered by all the professional departments of the College of Education. Graduate students are given liberal opportunities to carry on research projects under the direction of the faculty.

For further information address

DR. GEORGE F. ARPS
Dean, College of Education
Ohio State University
Columbus, Ohio

Of Interest to Teachers and Parents

New or recent books that explain modern educational thought and practice, or important phases thereof, and aid in solving the problems involved in the relations of home and school and community.

MOORE
PARENT, TEACHER, and SCHOOL—Price $2.00

This book is addressed to school administrators, teachers, and parents, with the purpose of making clearer to each group the problems of the others and with the hope of simplifying some of these problems. The author believes that the public school has become so closely connected with the home that it is highly desirable to supply informational literature that shall be of common interest to educators and parents.

GRUENBERG
OUTLINES OF CHILD STUDY—Price $1.80

"It is ... a useful service to offer a study of varied sources of information and opinion. The outlines and references of this book will prove helpful to many a troubled parent and teacher." *William H. Kilpatrick, Professor of Educational Psychology, Teachers College, Columbia University.*

KEITH AND BAGLEY
THE NATION AND THE SCHOOL—Price $1.60

This book outlines a solution for the problem of how the public schools can be made efficient on a nation-wide basis. The defects in our educational system are analyzed and traced to their sources and the two most serious weaknesses—poor rural schools and inadequate teacher preparation—are given special emphasis and attention.

Write for information.

THE MACMILLAN COMPANY

New York Chicago Dallas
Boston Atlanta San Francisco